INSECT HERBIVORES AS DRIVERS OF NATURAL SELECTION

Dissertation

zur

Erlangung der naturwissenschaftlichen Doktorwürde
(Dr. sc. nat.)

vorgelegt der

Mathematisch-naturwissenschaftlichen Fakultät

der

Universität Zürich

von

Tobias Züst

aus

Wolfhalden AR

Promotionskomitee:

Prof. Dr. Bernhard Schmid (Vorsitz)

PD Dr. Lindsay A. Turnbull (Leitung der Dissertation)

Prof. Dr. Daniel J. Kliebenstein

Prof. Dr. Kentaro Shimizu

Zürich, 2012

ARTIKEL 1:	General introduction
ARTIKEL 2:	A new method for measuring relative growth rate can uncover the costs of defensive compounds in *Arabidopsis thaliana* C. PAUL-VICTOR, T. ZÜST, M. REES, D. J. KLIEBENSTEIN & L. A. TURNBULL *New Phytologist*, 187, 1102-1111, 2010
ARTIKEL 3:	Using knockout mutants to reveal the growth costs of defensive traits T. ZÜST, B. JOSEPH, K. K. SHIMIZU, D. J. KLIEBENSTEIN & L. A. TURNBULL *Proceedings of the Royal Society B*, 278, 2598-2603, 2011
ARTIKEL 4:	Non-linear rosette area modelling for a simpler, more robust quantification of genotype effects T. ZÜST, C. PHILIPSON, A. HECTOR & L. A. TURNBULL *manuscript*
ARTIKEL 5	Quantification of growth rates and herbivory across RIL population of *Arabidopsis thaliana* reveals the costs of an effective pathway of plant defence T. ZÜST, M. IRAVANI, G. SCHAEPMAN-STRUB, J. A. CORWIN, B. LI, B. JOSEPH, D. J. KLIEBENSTEIN & L. A. TURNBULL *manuscript*
ARTIKEL 6	Natural enemies drive geographic variation in plant defenses T. ZÜST, C. HEICHINGER, U. GROSSNIKLAUS, R. HARRINGTON, D. J. KLIEBENSTEIN & L. A. TURNBULL *Science*, 338, 116-119, 2012
ARTIKEL 7	General Discussion

GENERAL INTRODUCTION

GENERAL INTRODUCTION

Diversity exists at all levels of biological organisation, ranging from widespread genetic diversity within populations of a single species to the almost bewildering variety of species found, for example, among rainforest trees. Biodiversity has not only intrinsic, but also clear economic value. For example, more diverse ecosystems tend to be more productive and more stable over time due to a larger potential to adapt to environmental change. Despite extensive research on the effects of biodiversity (e.g. Hooper *et al.*, 2005), the forces creating and maintaining biodiversity are still not fully understood. A number of theories have tried to explain the number of coexisting species in an ecosystem (e.g. Hutchinson, 1957; Hubbell, 2001). While classical niche-based theories can explain certain patterns of species distributions adequately, they often would predict fewer coexisting species than can be observed in natural systems. As a consequence, Hubbell (2001) proposed the neutral theory of biodiversity to explain species richness observed in such hyperdiverse systems, but the assumptions underlying this theory have been repeatedly shown to be unrealistic for natural communities of species (Gotelli & McGill, 2006; Purves & Turnbull, 2010).

A different set of theories employs various biotic interactions among organisms that create ‘biotic’ niches, for example selective pressures exerted by natural enemies such as herbivores, predators or pathogens (Janzen, 1970; Connell, 1971; Coley & Barone, 1996; Crawley, 1997; Grewell, 2008). Underlying the biotic niche concept are several assumptions: trade-offs prevent the evolution of all-powerful natural enemies and all-resistant hosts; the occurrence of a natural enemy is unpredictable in space or time; and resistance to a natural enemy trades-off with host fitness. If all assumptions are met, a defended host species will be favoured in the presence of natural enemies, but selected against in their absence. Host defences will be differentially effective against different natural enemies, thus a diverse community of natural enemies has the potential to maintain a large number of coexisting

species, where in their absence only the best competitor for abiotic resources would survive. Such fluctuating selective pressures at the landscape have been termed ‘selection mosaics’ (Thompson & Cunningham, 2002; Thompson, 2005), where change in locally favourable and unfavourable conditions maintains and creates biodiversity at the meta-population scale.

Even though the first two assumptions of the biotic niche concept are relatively undisputed, experimental evidence for this mechanism is still relatively rare (e.g. Gaines, 1985). Natural enemies such as herbivores are known to fluctuate strongly in space and time (Lawton & Gaston, 1989; Root & Cappuccino, 1992; Thrall & Burdon, 1997) and all-powerful enemies are rare and most herbivores specialise on a small number host plants. Similarly, most genetic and trait diversity within a species is usually associated with defence mechanisms (Frank, 1991; Bergelson *et al.*, 2001), which would be contradictory to the existence of all-resistant hosts. In contrast, the last prediction of the biotic niche concept has been challenged repeatedly. Even though defence traits are traditionally assumed to incur costs to the organism expressing them (Coley *et al.*, 1985; Bazzaz *et al.*, 1987; Herms & Mattson, 1992), experimental studies often struggled to find the expected trade-offs with fitness, leading several studies to conclude that at least some defence traits might not be costly to produce (reviewed in Koricheva, 2002; Koricheva *et al.*, 2004).

Natural defences: useful but costly

Plants employ a range of defence mechanisms to deter and avoid herbivory. Physical structures such as thorns and spines impede herbivore feeding directly and chemical defences deter herbivores once they begin feeding on a plant. Any such trait is generally only effective against a subset of the herbivore community, and specialisation for defended host plants is

common among herbivore species, for example through the evolution of detoxification mechanisms (Ratzka *et al.*, 2002; Wittstock *et al.*, 2004; Kazana *et al.*, 2007).

Three different types of defence costs have been identified (Koricheva, 2002): allocation costs, (Rhoades, 1979; Bazzaz *et al.*, 1987) which are trade-offs between defence, growth and reproduction within an individual plant; opportunity costs (Coley *et al.*, 1985), which are a special form of allocation costs that are affected by ontogeny; and ecological costs (Simms, 1992; Rausher, 1996), which are costs that result when the expression of a defence trait makes a host susceptible to other herbivores or has deleterious effects on pollinators, predators or parasitoids. Not all of these costs are easily measurable, and failure to account for the different types has been claimed to be responsible for the inconsistent detection of defence costs (Koricheva, 2002). While this might be true for many plant species, it is unlikely to be the case for annual plants. Annual plants only have a short lifetime, are generally fast growing and are highly resource limited. Under such conditions, even small increases in nutrient availability can leads to marked changes in plant growth (e.g. Paul-Victor *et al.*, 2010). Any allocation of resources away from growth to defence traits should therefore become clearly apparent in annual plants, and hence allocation costs should be easily detectable.

Costs are measured as the reduction of plant fitness caused by the expression of a trait. While the seed set of a plant is often considered to be the best estimate of lifetime fitness, plant growth rate can also be a good predictor of the outcome of competition and is thus a useful surrogate for fitness in competitive environments (Fakheran *et al.*, 2010; Züst *et al.*, 2011).

Describing growth rate

Growth rate of a plant is traditionally described as the relative growth rate (RGR). RGR is typically based on only two biomass measurements and is calculated using:

$$RGR = \frac{log(M_2/M_1)}{t_2 - t_1}$$

The advantage of such calculations is that they are easy to carry out and many species or genotypes can be compared for relatively little time and effort (Hunt, 1982). However, this formulation implicitly assumes that plants grow exponentially (linear on the log-scale), in which case RGR would be constant and unbiased by size. In fact, as plants get larger, deviations from exponential growth occur for a variety of reasons, including increasing allocation to structural non-photosynthetic tissue, self-shading of the leaf canopy and increasing below-ground resource limitation (Evans, 1972; Ingestad & Agren, 1992; Maranon & Grubb, 1993; Enquist *et al.*, 1999). In addition, RGR is itself size-dependent and declines as individual plants grow (Hunt, 1982; Hunt & Cornelissen, 1997; Enquist *et al.*, 1999; West *et al.*, 2001); hence larger individuals are expected to have lower RGR than smaller individuals when measured over the same time period. This can confound analyses when the species or genotypes differ in their initial sizes (Turnbull *et al.*, 2008; Rose *et al.*, 2009).

The solution to these problems is the use of non-linear growth functions that assume biologically more reasonable growth processes, and to calculate size-standardized RGR or SGR that facilitate comparison among species or genotypes at a common reference size (Rose *et al.*, 2009; Paul-Victor *et al.*, 2010; Paine *et al.*, 2011). When examining growth/defence trade-offs, a size-standardised analysis should thus reveal whether each new unit of defended tissue is more costly to make than each new unit of undefended tissue for plants of standardised size.

Arabidopsis thaliana as a model species

The thale cress *Arabidopsis thaliana* L. Heynh (Brassicaceae) has a long history as a model plant in biology. Almost 70 years ago, Laibach (1943) proposed the use of *A. thaliana* as a model plant, as its short generation time, high fecundity, and the ease with which plants could be crossed and mutants could be generated provides large potential for targeted study of plant traits. *A. thaliana* is a highly and readily selfing plant, and outcrossing rates in natural populations are estimated in the range of 0.3 to 2.5 % (Abbott & Gomes, 1989; Bergelson *et al.*, 1998; Bakker *et al.*, 2006), although it can be much higher in some populations (e.g. Bomblies *et al.*, 2010). *A. thaliana* is highly variable in many traits, and a large part of this variability is represented in natural variants (accessions) that have been collected from the natural habitats of the species worldwide and are available to researchers from stock centres. In addition, there are thousands of mutants in which specific genes have been artificially modified and which allow targeted study of the importance of specific genes. *A. thaliana* produces leaf hairs (trichomes) which cover leaves and stem and have been shown to reduce feeding damage by natural herbivore assemblages (Mauricio, 1998). *A. thaliana* also produces chemical defences, mainly secondary metabolites belonging to the group of glucosinolates, which are present in all parts of the plant (Vaughan *et al.*, 1976). Both groups of traits are assumed to be costly, as both traits are negatively correlated with the seed set (Mauricio, 1998) and the growth rate of plants (Paul-Victor *et al.*, 2010; Züst *et al.*, 2011).

Glucosinolates are produced by all Brassicaceae and have known defensive properties against herbivory (Bones & Rossiter, 1996). Over 120 different glucosinolate compounds have been identified in crucifer species (Benderoth *et al.*, 2006), with at least 43 different glucosinolate compounds present in *A. thaliana* (Reichelt *et al.*, 2002; Kliebenstein *et al.*, 2007). A glucosinolate consist of a sulphur-linked glycone moiety, a nitrogen-linked sulphate, and a variable side chain (Mithen *et al.*, 1995). This side-chain is the biological active part

and may contain aliphatic, indolyl, or aromatic groups. In order to become toxic, glucosinolates require enzymatic activation by enzymes, generally by tissue disruption and subsequent combination of the spatially separate enzymes and glucosinolate compounds (Kelly *et al.*, 1998). Variation in aliphatic glucosinolates among natural accessions of *A. thaliana* has been explained by the combination of functional and non-functional (null) alleles at four major loci. These are *GS-ELONG*, which regulates the carbon side-chain elongation of aliphatic glucosinolates; *GS-AOP*, which controls the conversion of methylsulfinylalkyl to either alkenyl (*ALK*) or hydroxypropyl (*OHP*) glucosinolates; *GS-OH*, responsible for the conversion of 3-butenyl to 2-hydroxy-3-butenyl glucosinolate; and *GS-OX*, which regulates the conversion of methylthioalkyl to methylsulfinylalkyl glucosinolates (Mithen *et al.*, 1995; Kliebenstein *et al.*, 2001). The combination of heritable polymorphisms at these four loci leads to the formation of 'chemotypes', i.e., a group of genotypes that accumulate a characteristic set of glucosinolates. The chemotype of *A. thaliana* plays an important role in determining the herbivore community that feeds on a given chemotype. For example, high levels of alkenyl glucosinolates are generally associated with reduced damage caused by generalist caterpillars (Lambrix *et al.*, 2001; Kliebenstein *et al.*, 2002; Kliebenstein, 2004). However, the same glucosinolates act as stimulants for feeding and oviposition by numerous specialists, such as cabbage aphids, cabbage seed weevils, cabagge root flies, turnip aphids, and diamondback moths (Raybould & Moyes, 2001). In contrast, Bidart-Bouzat and Kliebenstein (2008) carried out an herbivory assay in the field with a predominantly specialist herbivore community, and found the least levels of damage associated with *3C-OHP* and *4C-NULL* chemotypes. This is a clear indication that a selection mosaic of herbivores could be responsible for the large genetic and chemical diversity found in natural populations of *A. thaliana.*

Goal of the project

Trade-offs of between defence traits and fitness are generally assumed to exist but are not consistently found in experimental studies. We therefore asked the question whether failure to detect such trade-offs can be caused by methodological problems with fitness surrogates (Chapter 1). We measured both conventional growth rates (RGRs) and size-standardised growth rates (SGRs) on a subset of an *A. thaliana* recombinant inbred line (RIL) population and related these measures to levels of defensive compounds. Such correlations of two traits never allow causal conclusions on their relation, and the expression of any trait is usually confounded with other non-target traits. As a solution to this problem, we used knockout mutants of *A. thaliana* to measure the associated allocation costs of these traits (Chapter 2). Knockout mutants ideally differ only in one trait of interest, and hence are a useful method to directly compare the performance of otherwise genetically identical plants expressing or lacking a trait.

SGR is a better method to describe growth as it often produces biologically more meaningful results, but it also has the disadvantage of requiring a higher sampling effort. High-throughput methods of growth phenotyping are becoming available to deal with this problem, but often still use inadequate statistical methods to process the data they produce. To demonstrate this, we have taken the data from a recently published paper and analysed it with a much simpler, more meaningful method than the method chosen by the authors of the study (Chapter 3). In addition, we have developed a cheap, high-accuracy and high-throughput method to measure the rosette growth of *A. thaliana* and applied it to a new RIL population in search of fitness/defence trade-offs (Chapter 4).

Trade-offs measured on individual plants in a single generation can never adequately describe costs, since the real fitness of a plant is often affected by many more variables, which are not accounted for in such studies. A better way to measure fitness costs is to expose plants

under near-natural conditions to a selective pressure and observe the proportion of offspring in future generation. Combining this approach with a direct test of the biotic niche theory and the selection mosaic concept, we carried out a selection experiment in which we exposed a community of *A. thaliana* accessions for several generations to various constant herbivore pressures (Chapter 5), in this way testing all the major assumptions of the biotic niche concept.

All the chapters are presented in manuscript format and already published chapters are presented in their final published version.

REFERENCES

Abbott RJ, Gomes MF. 1989. Population genetic structure and outcrossing rate of *Arabidopsis thaliana* (L.) Heynh. *Heredity* **62**: 411-418.

Bakker EG, Stahl EA, Toomajian C, Nordborg M, Kreitman M, Bergelson J. 2006. Distribution of genetic variation within and among local populations of *Arabidopsis thaliana* over its species range. *Molecular Ecology* **15**: 1405-1418.

Bazzaz FA, Chiariello NR, Coley PD, Pitelka LF. 1987. Allocating resources to reproduction and defense. *Bioscience* **37**: 58-67.

Benderoth M, Textor S, Windsor AJ, Mitchell-Olds T, Gershenzon J, Kroymann J. 2006. Positive selection driving diversification in plant secondary metabolism. *Proceedings of the National Academy of Sciences of the United States of America* **103**: 9118-9123.

Bergelson J, Kreitman M, Stahl EA, Tian DC. 2001. Evolutionary dynamics of plant R-genes. *Science* **292**: 2281-2285.

Bergelson J, Stahl E, Dudek S, Kreitman M. 1998. Genetic variation within and among populations of *Arabidopsis thaliana*. *Genetics* **148**: 1311-1323.

Bidart-Bouzat MG, Kliebenstein DJ. 2008. Differential levels of insect herbivory in the field associated with genotypic variation in glucosinolates in *Arabidopsis thaliana*. *Journal of Chemical Ecology* **34**: 1026-1037.

Bomblies K, Yant L, Laitinen RA, Kim S-T, Hollister JD, Warthmann N, Fitz J, Weigel D. 2010. Local-scale patterns of genetic variability, outcrossing, and spatial structure in natural stands of *Arabidopsis thaliana*. *PLoS Genetics* **6**.

Bones AM, Rossiter JT. 1996. The myrosinase-glucosinolate system, its organisation and biochemistry. *Physiologia Plantarum* **97**: 194-208.

Coley PD, Barone JA. 1996. Herbivory and plant defenses in tropical forests. *Annual Review of Ecology and Systematics* **27**: 305-335.

Coley PD, Bryant JP, Chapin FS. 1985. Resource availability and plant antiherbivore defense. *Science* **230**: 895-899.

Connell JH 1971. On the role of natural enemies in preventing competitive exclusion in some marine animals and in rainforest trees. In: P. J. den Boer and G. R. Gradwell eds. *Dynamics of Populations*. Wageningen: Center for Agricultural Publishing and Documentation, 298-312.

Crawley MJ 1997. Plant-herbivore dynamics. In: M. J. Crawley ed. *Plant Ecology*. London: Blackwell Science Ltd, 401-474.

Enquist BJ, West GB, Charnov EL, Brown JH. 1999. Allometric scaling of production and life-history variation in vascular plants. *Nature* **401**: 907-911.

Evans GC. 1972. *The quantitative analysis of plant growth*. Oxford: Blackwell Scientific Publications.

Fakheran S, Paul-Victor C, Heichinger C, Schmid B, Grossniklaus U, Turnbull LA. 2010. Adaptation and extinction in experimentally fragmented landscapes. *Proceedings of the National Academy of Sciences of the United States of America* **107**: 19120-19125.

Frank SA. 1991. Ecological and genetic models of host pathogen coevolution. *Heredity* **67**: 73-83.

Gaines SD. 1985. Herbivory and between-habitat diversity: the differential effectiveness of defenses in a marine plant. *Ecology* **66**: 473-485.

Gotelli NJ, McGill BJ. 2006. Null versus neutral models: what's the difference? *Ecography* **29**: 793-800.

Grewell BJ. 2008. Parasite facilitates plant species coexistence in a coastal wetland. *Ecology* **89**: 1481-1488.

Herms DA, Mattson WJ. 1992. The dilemma of plants: to grow or defend. *Quarterly Review of Biology* **67**: 283-335.

Hooper DU, Chapin FS, Ewel JJ, Hector A, Inchausti P, Lavorel S, Lawton JH, Lodge DM, Loreau M, Naeem S, Schmid B, Setala H, Symstad AJ, Vandermeer J, Wardle DA. 2005. Effects of biodiversity on ecosystem functioning: A consensus of current knowledge. *Ecological Monographs* **75**: 3-35.

Hubbell SP. 2001. *The unified neutral theory of biodiversity and biogeography*. Princeton, NJ, USA: Princeton University Press.

Hunt R. 1982. *Plant growth curves: a functional approach to plant growth analysis*. London: Edward Arnold.

Hunt R, Cornelissen JHC. 1997. Physiology, allocation, and growth rate: A reexamination of the Tilman model. *American Naturalist* **150**: 122-130.

Hutchinson GE. 1957. Population studies - Animal ecology and demography - Concluding remarks. *Cold Spring Harbor Symposia on Quantitative Biology* **22**: 415-427.

Ingestad T, Agren GI. 1992. Theories and methods on plant nutrition and growth. *Physiologia Plantarum* **84**: 177-184.

Janzen DH. 1970. Herbivores and number of tree species in tropical forests. *American Naturalist* **104**: 501-&.

Kazana E, Pope TW, Tibbles L, Bridges M, Pickett JA, Bones AM, Powell G, Rossiter JT. 2007. The cabbage aphid: a walking mustard oil bomb. *Proceedings of the Royal Society B: Biological Sciences* **274**: 2271-2277.

Kelly PJ, Bones A, Rossiter JT. 1998. Sub-cellular immunolocalization of the glucosinolate sinigrin in seedlings of *Brassica juncea*. *Planta* **206**: 370-377.

Kliebenstein DJ. 2004. Secondary metabolites and plant/environment interactions: a view through *Arabidopsis thaliana* tinged glasses. *Plant Cell and Environment* **27**: 675-684.

Kliebenstein DJ, D'Auria JC, Behere AS, Kim JH, Gunderson KL, Breen JN, Lee G, Gershenzon J, Last RL, Jander G. 2007. Characterization of seed-specific

benzoyloxyglucosinolate mutations in *Arabidopsis thaliana*. *Plant Journal* **51**: 1062-1076.

Kliebenstein DJ, Kroymann J, Brown P, Figuth A, Pedersen D, Gershenzon J, Mitchell-Olds T. 2001. Genetic control of natural variation in *Arabidopsis* glucosinolate accumulation. *Plant Physiology* **126**: 811-825.

Kliebenstein DJ, Pedersen D, Barker B, Mitchell-Olds T. 2002. Comparative analysis of quantitative trait loci controlling glucosinolates, myrosinase and insect resistance in *Arabidopsis thaliana*. *Genetics* **161**: 325-332.

Koricheva J. 2002. Meta-analysis of sources of variation in fitness costs of plant antiherbivore defenses. *Ecology* **83**: 176-190.

Koricheva J, Nykanen H, Gianoli E. 2004. Meta-analysis of trade-offs among plant antiherbivore defenses: Are plants jacks-of-all-trades, masters of all? *American Naturalist* **163**: E64-E75.

Laibach F. 1943. *Arabidopsis thaliana* (L.) Heynh. als Objekt für genetische und entwicklungsphysiologische Untersuchungen. *Botanisches Archiv* **44**: 439-455.

Lambrix V, Reichelt M, Mitchell-Olds T, Kliebenstein DJ, Gershenzon J. 2001. The Arabidopsis epithiospecifier protein promotes the hydrolysis of glucosinolates to nitriles and influences *Trichoplusia ni* herbivory. *Plant Cell* **13**: 2793-2807.

Lawton JH, Gaston KJ. 1989. Temporal patterns in the herbivorous insects of bracken: a test of community predictability. *Journal of Animal Ecology* **58**: 1021-1034.

Maranon T, Grubb PJ. 1993. Physiological-basis and ecological significance of the seed size and relative growth-rate relationship in mediterranean annuals. *Functional Ecology* **7**: 591-599.

Mauricio R. 1998. Costs of resistance to natural enemies in field populations of the annual plant *Arabidopsis thaliana*. *American Naturalist* **151**: 20-28.

Mithen R, Clarke J, Lister C, Dean C. 1995. Genetics of aliphatic glucosinolates. III. Side-chain structure of aliphatic glucosinolates in *Arabidopsis thaliana*. *Heredity* **74**: 210-215.

Paine CET, Marthews TR, Vogt DR, Purves D, Rees M, Hector A, Turnbull LA. 2011. How to fit nonlinear plant growth models and calculate growth rates: an update for ecologists. *Methods in Ecology and Evolution*.

Paul-Victor C, Züst T, Rees M, Kliebenstein DJ, Turnbull LA. 2010. A new method for measuring relative growth rate can uncover the costs of defensive compounds in *Arabidopsis thaliana*. *New Phytologist* **187**: 1102-1111.

Purves DW, Turnbull LA. 2010. Different but equal: the implausible assumption at the heart of neutral theory. *Journal of Animal Ecology* **79**: 1215-1225.

Ratzka A, Vogel H, Kliebenstein DJ, Mitchell-Olds T, Kroymann J. 2002. Disarming the mustard oil bomb. *Proceedings of the National Academy of Sciences of the United States of America* **99**: 11223-11228.

Rausher MD. 1996. Genetic analysis of coevolution between plants and their natural enemies. *Trends in Genetics* **12**: 212-217.

Raybould AF, Moyes CL. 2001. The ecological genetics of aliphatic glucosinolates. *Heredity* **87**: 383-391.

Reichelt M, Brown PD, Schneider B, Oldham NJ, Stauber E, Tokuhisa J, Kliebenstein DJ, Mitchell-Olds T, Gershenzon J. 2002. Benzoic acid glucosinolate esters and other glucosinolates from *Arabidopsis thaliana*. *Phytochemistry* **59**: 663-671.

Rhoades DF 1979. Evolution of plant chemical defense against herbivores. In: G. A. Rosenthal and D. H. Janzen eds. *Herbivores: their interaction with secondary metabolites*. New York: Academic Press, 3-54.

Root RB, Cappuccino N. 1992. Patterns in population change and the organization of the insect community associated with goldenrod. *Ecological Monographs* **62**: 393-420.

Rose KE, Atkinson RL, Turnbull LA, Rees M. 2009. The costs and benefits of fast living. *Ecology Letters* **12**: 1379-1384.

Simms EL 1992. Costs of plant resistance to herbivory. In: R. S. Fritz and E. L. Simms eds. *Plant resistance to herbivores and pathogens: ecology, evolution and genetics.* Chicago, Illinois, USA: University of Chicago Press, 392-425.

Thompson JN. 2005. *The geographic mosaic of coevolution*. Chicago: University of Chicago Press.

Thompson JN, Cunningham BM. 2002. Geographic structure and dynamics of coevolutionary selection. *Nature* **417**: 735-738.

Thrall PH, Burdon JJ. 1997. Host-pathogen dynamics in a metapopulation context: the ecological and evolutionary consequences of being spatial. *Journal of Ecology* **85**: 743-753.

Turnbull LA, Paul-Victor C, Schmid B, Purves DW. 2008. Growth rates, seed size, and physiology: Do small-seeded species really grow faster? *Ecology* **89**: 1352-1363.

Vaughan JG, Macleod AJ, Jones BMG. 1976. *The biology and chemistry of the Cruciferae*. London: Academic Press.

West GB, Brown JH, Enquist BJ. 2001. A general model for ontogenetic growth. *Nature* **413**: 628-631.

Wittstock U, Agerbirk N, Stauber EJ, Olsen CE, Hippler M, Mitchell-Olds T, Gershenzon J, Vogel H. 2004. Successful herbivore attack due to metabolic diversion of a plant chemical defense. *Proceedings of the National Academy of Sciences of the United States of America* **101**: 4859-4864.

Züst T, Joseph B, Shimizu KK, Kliebenstein DJ, Turnbull LA. 2011. Using knockout mutants to reveal the costs of defensive traits. *Proceedings of the Royal Society B: Biological Sciences* **278**: 2598-2603.

New Phytologist

A new method for measuring RGR can uncover the costs of defensive compounds in *Arabidopsis thaliana*.

Cloé Paul-Victor[1], Tobias Züst[1], Mark Rees[2], Daniel J. Kliebenstein[3] and Lindsay A. Turnbull[1].

[1] Institute of Evolutionary Biology and Environmental Studies, Winterthurerstrasse 190, University of Zurich, Zurich, Switzerland, CH-8057; [2] Department of Animal and Plant Sciences, University of Sheffield, Western Bank, Sheffield S10 2TN, UK; [3] Department of Plant Sciences, University of California, Davis, California 95616, USA.

Author for correspondence:
Lindsay A. Turnbull
Tel: 0041 44 635 6120
Fax: 0041 44 635 5711
Email: lindsay.turnbull@ieu.uzh.ch

Received: 15 March 2010
Accepted: 1 May 2010

Summary

• Most plants suffer some degree of herbivore attack and many actively defend themselves against such an event. However, while such defence is generally assumed to be costly, it has sometimes proved difficult to demonstrate the costs of defensive compounds.

• Here, we present a method for analysing growth rates which allows the effects of variation in initial plant size to be properly accounted for and apply it to 30 lines from a recombinant inbred (RIL) population of *Arabidopsis thaliana*. We then relate different measures of relative growth rate (RGR) to damage caused by a specialist lepidopteran insect and to levels of putative defensive compounds measured on the same lines.

• We show that seed size variation within the RIL population is large enough to generate differences in RGR, even when no other physiological differences exist. However, once size-standardised, RGR was positively correlated with herbivore damage (fast-growing lines suffered more damage) and was negatively correlated with the concentration of several glucosinolate compounds.

• We conclude that defensive compounds do have a growth cost and that the production of such compounds results in reduced herbivore damage. However, size standardisation of RGR was essential to uncovering the growth costs of defensive compounds.

Key Words: RGR, trade-off, herbivore, defence, *Arabidopsis thaliana*, glucosinolate.

New Phytologist (2010) final accepted version

Introduction

Plants differ in their growth rates and some of this variation is thought to be negatively correlated with traits such as herbivore defences (Bazzaz *et al.*, 1987; Herms & Mattson, 1992), frost resistance (Agrawal *et al.*, 2004; Turnbull *et al.*, 2008) or investment in non-photosynthetic structures such as storage organs (Poorter & Kitajima, 2007). Trade-offs – or negative correlations – between growth rates and investment in defence are predicted by life-history theory because defence is assumed to be costly (Bazzaz *et al.*, 1987; Perrin & Sibly, 1993; Iwasa, 2000). A plant that does not invest in defence can therefore grow more rapidly but it should suffer more damage when herbivores strike; conversely, if a plant invests in defensive compounds it should grow more slowly but suffer less damage. Such trade-offs are usually assumed in theoretical work (Perrin & Sibly, 1993; Iwasa, 2000) but, experimental demonstrations have sometimes proved more difficult (reviewed in Koricheva, 2002), particularly for compounds associated with defence against herbivores (Bergelson & Purrington, 1996; Arendt, 1997; Arendt, 2000); (Almeida-Cortez *et al.*, 1999; Almeida-Cortez & Shipley, 2002; Siemens *et al.*, 2002). Here we explore whether the methods commonly used to calculate growth rates, which fail to account for differences in initial size (Hunt, 1982; Hunt & Cornelissen, 1997), are partly responsible for the difficulties in detecting negative correlations between growth rates and defence, as outlined below.

The problem with RGR

The most widely-used method to compare growth rates among species or genotypes is relative growth rate (RGR)

$$RGR = \frac{\log(M_2 / M_1)}{t_2 - t_1} \qquad \text{eqn 1}$$

where M_i is the mass of the plant at time t_i. Experiments using such calculations are easy to carry out and many

species or genotypes can be compared for relatively little time and effort (Hunt, 1982). The problem with such calculations is that RGR is itself size-dependent and declines as individual plants grow (Hunt, 1982; Hunt & Cornelissen, 1997; Enquist *et al.*, 1999; West *et al.*, 2001); hence larger individuals are expected to have lower RGR than smaller individuals when measured over the same time period. This can confound analyses when the species or genotypes differ in their initial sizes (Turnbull *et al.*, 2008; Rose *et al.*, 2009). To overcome this problem, we need to carry out a size-standardised analysis in which species are compared at a common size. When examining growth/defence trade-offs, a size-standardised analysis should reveal whether each new unit of defended tissue is more costly to make than each new unit of undefended tissue for plants of standardised size.

Defence in *Arabidopsis thaliana*

Arabidopsis thaliana produces a variety of secondary metabolites associated with defence. This chemical arsenal consists of a group of glucosinolates, alongside protease inhibitors, phenolics and terpenoid volatiles (Kliebenstein, 2004). Glucosinolates are amino-acid derived thioglycosides consisting of a conserved core structure and a highly diverse side chain. So far, at least 43 different glucosinolate compounds have been identified in *Arabidopsis* (Reichelt *et al.*, 2002; Kliebenstein *et al.*, 2007), the majority of which have an aliphatic side-chain, while another group of glucosinolates has indolic side-chains (Kliebenstein *et al.*, 2001b). Glucosinolates serve as a major chemical defence mechanism against insect herbivores, bacteria and fungi (Bones & Rossiter, 1996). For example, a number of studies have indicated that high glucosinolate content can delay larval development and reduce the survival of leaf-chewing lepidopteran insects (Kliebenstein *et al.*, 2002; Barth & Jander, 2006; Beekwilder *et al.*, 2008). Different types of herbivores are also affected by different glucosinolate compounds, for example, phloem-feeding aphids are mainly impaired by indolic glucosinolates (Kim & Jander, 2007).

Given their molecular structure, accumulation of glucosinolates by *Arabidopsis* might be expected to incur some metabolic or regulatory cost, leading to reductions in growth rate. However, when looked for, such growth costs have not been detected (e.g. Siemens *et al.*, 2002). Given that *Arabidopsis* lines vary in seed size and emergence time (germination day), comparisons among lines carried out over a fixed time period inevitably compare lines at different sizes. The failure to detect the growth costs of glucosinolates could therefore be due to the lack of size-standardisation when calculating growth rates.

Here we present a method for calculating size-standardised RGR which requires multiple harvests and apply it to data collected on 30 lines of *Arabidopsis* from a recombinant inbred (RIL) population. We then combine this growth data with published data on the same RIL population to examine the correlations between 1) growth rates and the concentrations of several glucosinolate compounds and 2) growth rates and herbivore damage inflicted by a specialist insect.

MATERIAL AND METHODS

Plant material

To demonstrate the potentially confounding effects of seed size on growth rates, we selected a RIL population derived from crosses between two accessions of *Arabidopsis thaliana*: the small-seeded Landsberg *erecta* (L*er*: mean mass of 100 seeds ± 1 SD: 1.93 mg ± 0.10) and the large-seeded Cape Verde Islands (Cvi: mean mass of 100 seeds ± 1 SD: 3.51 mg ± 0.08) (Alonso-Blanco *et al.*, 1998; Alonso-Blanco *et al.*, 1999). For the growth experiment, we selected 30 RILs from the possible set of 162. The 30 lines were selected by dividing the original 162 lines into six equally-spaced seed mass groups and selecting five lines at random from each group. Half of the selected lines carry the *erecta* mutation inherited from the L*er* parent, while the other half carries the wild-type *ERECTA* allele (Table S1). Lines carrying the *erecta* mutation have reduced height and different flower morphologies (Arabidopsis Biological Resource Centre (ABRC)). A summary of published information about the lines is available in Table S1. The seeds were obtained from The Arabidopsis Information Resource (TAIR) and we estimated sown seed mass by weighing one batch of 100 seeds from each of the 30 selected lines.

Experimental design

Plants were grown in small (20 mm diameter), medium

Table 1 Schedule of harvest dates showing the average developmental stage observed at each harvest. On average, germination occurred 4.7 days after sowing.

Harvest	Days after sowing	Average age (Days after germination)	Developmental stage
1	7	2.3	2 leaves
2	11	6.3	4 leaves
3	15	10.3	6 leaves
4	20	15.3	8 leaves and bolting
5	28	23.3	First flowers seen
6	33	28.3	First fruits seen

(30 mm diameter) and large cylinders (40 mm diameter) inserted into standardized cells (65 mm diameter) within a flat completely filled with a mixture of 50% sand and 50% compost. Each flat contained 35 cells and was 70 mm deep. The cylinders allowed us to randomise pot diameter treatments within flats and ensured that the spacing of individuals in different pot sizes and the surface area available to growing rosettes was exactly the same. However, the three pot sizes provide different degrees of belowground growth restriction (Paul-Victor & Turnbull, 2009). Pots were sown with four seeds and thinned as soon as seedlings emerged to leave one plant per pot (the most central healthy seedling). The plants were grown in a glasshouse with both natural light and additional artificial lighting which came on automatically when the natural light was below 25 kLux and kept under a cycle of 16 h light (22°C) and 8 h dark (20°C). Germination, bolting (initiation of the flowering stem) and flowering (opening of the first flower) were recorded for each plant to the nearest day.

Biomass was collected during six sequential, destructive harvests. We separated the plant parts into roots, rosette leaves and inflorescence (when present) and counted the number of leaves. Plant parts were dried at 80°C for three days and weighed to the nearest microgram. We focussed on the active stages of plant growth by harvesting at relevant points of the plants' development; thus each harvest represents a developmental stage observed in most individuals (Fig. 1 and Table 1). By the last harvest (33 days after sowing) no siliques were observed to have opened and hence no biomass was lost as seeds; however, rosette growth had mostly stopped (evidenced by relatively little change in rosette mass between harvests 4 and 5). At each harvest there were two replicates of each line and pot size combination, giving 1080 plants in total. A few plants are missing due to germination failures in the growth experiment.

Size-standardised RGR

We modelled total biomass (rosette + roots + inflorescence) as a function of plant age (days since germination) using a three-parameter asymptotic regression model. Plant biomass was log-transformed giving:

$$\log\left(M_{i,t}\right) = A_i + (\log(M_{i,0}) - A_i)\exp(-\exp(r_i)t) \qquad \text{eqn 2}$$

where $M_{i,0}$ is the starting mass at $t = 0$, A_i is the asymptotic mass as $t \to \infty$ and r_i is the logarithm of the rate constant (the rate constant is log-transformed to ensure positive growth). The time required to reach a given a reference mass, M_{ref}, is given by

$$t(M_{ref}) = \log\left(\frac{\log(M_{i,0}) - A_i}{\log(M_{ref}) - A_i}\right) \times \exp(-r_i) \qquad \text{eqn 3}$$

RGR is given by $d(\log(M_t))/dt$, hence we can calculate size-standardised RGR by differentiating eqn 2 and substituting for $t=t(M_{ref})$. This gives

$$RGR_i = \exp(r_i)\left(A_i - \log(M_{ref})\right) \qquad \text{eqn 4}$$

Thus, size-standardised RGR declines with mass and depends on three parameters, the rate constant (r_i) the asymptotic mass (A_i) and the reference mass (M_{ref}).

To calculate size-standardised RGR for each of the 30 lines we fitted the above model using the function *nlme* in the statistical package R (R Development Core Team, 2008). Lines were treated as a random effect and pot volume and seed mass as fixed effects. Throughout, we followed the model-building approach advocated by the developers of *nlme* (Pinheiro & Bates, 2000) which includes assessment and removal of non-significant terms. The significance of fixed effects (pot volume and seed mass) was assessed using F-tests while the significance of the random effects (lines) was assessed using likelihood ratio tests (Pinheiro & Bates, 2000).

Table 2 Estimates of the fixed effects from the final growth model. Random, i.e. line effects were retained for r_i.

	Estimate	S.E.	t-value	p-value
Asymptotic mass (Asym)	2.41	0.0785	30.7	<.0001
Asym (Pot diameter = 30)	1.32	0.117	11.2	<.0001
Asym (Pot diameter = 40)	1.84	0.128	14.4	<.0001
Rate parameter (r_i)	-2.20	0.0493	-44.7	<.0001
r_i (Pot diameter = 30)	-0.204	0.0507	-4.02	0.0001
r_i (Pot diameter = 40)	-0.296	0.0509	-5.82	<.0001
$M_{i,0}$ (intercept)	-0.097	0.660	-0.147	0.883
$M_{i,0}$ (log(sown.seed.mass))	0.753	0.182	4.14	<.0001

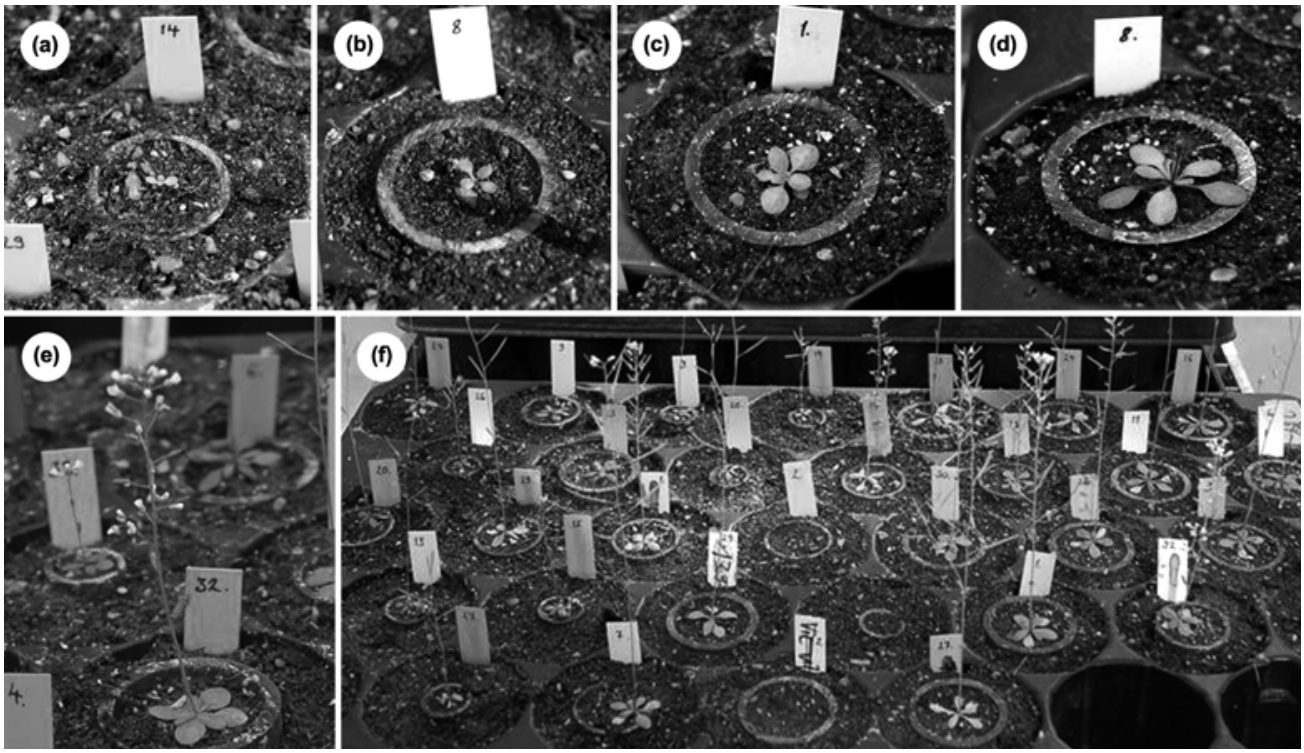


Figure 1 Picture of the experiment showing the developmental stages of the plants at each of the six harvests (pictures a–f correspond to harvests 1–6; see Table 1). Note that the surface area available to grow rosettes is exactly the same for the three pot size treatments.

Conventional RGR

Conventional RGR is an average over some specified time period. Average RGR (RGR_{av}) is typically measured by making two harvests separated by a short time interval and applying eqn 1. Here we calculate average RGR_{av} over the whole growth interval (harvest 1 – harvest 6). We also calculated early RGR (RGR_{early}) using data from the first two harvests (conducted 7 and 11 days after sowing).

Secondary compounds and herbivory

Estimates of growth rates (RGR_{av}, RGR_{early} and size-standardised RGR) in the largest pot size (diameter = 40 mm) were used to test associations between different measures of growth rate with herbivore damage and with glucosinolate concentrations in leaves and seeds. In an earlier experiment, Kliebenstein et al. (2001a) measured glucosinolate concentration in leaves and seeds on the same RILs and recorded levels of damage inflicted by two insects after feeding for a short time interval (Kliebenstein *et al.*, 2002). Although glucosinolate concentrations, damage by herbivores and growth rates were not measured in the same individuals, the genetic stability of a RIL population allows data from different experiments to be compared as long as strong environment x genotype interactions are lacking (West *et al.*, 2006; Keurentjes *et al.*, 2007; Sønderby *et al.*, 2007; Wentzell *et al.*, 2007; Hansen *et al.*, 2008; Keurentjes *et al.*, 2008; Fu *et al.*, 2009; Sulpice *et al.*, 2009).

For glucosinolate content, five plants per RIL were planted individually in separate pots (diameter = 60 mm), replicated three times. After three weeks, 10 leaves were harvested from each replicate while plants were grown to senescence and seeds collected within each replicate. Both leaf and seed samples were extracted and analysed with previously-described high-throughput methods (Kliebenstein *et al.*, 2001a; Kliebenstein *et al.*, 2001b). Herbivory assays using two lepidopteran species were conducted in another experiment (Kliebenstein *et al.*, 2002). A single 1st instar larva of the specialist *Plutella xylostella* L. or the generalist *Trichoplusia ni* Hübner were placed on 4-week-old plants and the area removed by the herbivores after 48 hours of feeding was measured. Herbivory estimates in *Arabidopsis* based on leaf area removal are highly correlated with herbivory estimates based on larval weight gain (Jander *et al.*, 2001; Barth & Jander, 2006), confirming the reliability of this method. Each RIL was assayed for damage by each lepidopteran species 16 independent times. As *P. xylostella* removed large proportions of the rosettes it was necessary to correct the herbivory scores for rosette size. This was done by fitting a randomized complete blocks ANOVA using the model HERBIVORY = CONSTANT + FLAT + LINE + SIZE, where flat is a blocking term. The predicted herbivory means were then taken as size-standardised herbivory scores. Previous analyses reveal that there is significant among-line variation in both damage inflicted by herbivores and in the concentrations of various glucosinolates compounds (Kliebenstein *et al.*, 2002).

Table 3. Correlations (Pearson's product-moment) between glucosinolate compounds in leaves and seeds and the three different measures of RGR. Significant correlations are highlighted in boldface. The concentrations of some glucosinolate compounds were not measured in the leaves; hence this correlation is not available (NA).

		RGR_{av}		RGR_{early}		Size-standardised RGR	
indolyl-3-methyl	leaves	0.276	*p* = 0.141	0.217	*p* = 0.268	-0.228	*p* = 0.226
	seed	0.291	*p* = 0.119	-0.350	*p* = 0.068	-0.226	*p* = 0.231
1-methoxy-indolyl-3-methyl	leaves	NA	NA	NA	NA	NA	NA
	seed	0.211	*p* = 0.263	0.102	*p* = 0.607	-0.336	*p* = 0.070
4-methoxy-indolyl-3-methyl	leaves	0.113	*p* = 0.552	0.316	*p* = 0102	-0.102	*p* = 0.593
	seed	NA	NA	NA	NA	NA	NA
Total indolic glucosinolates	leaves	0.314	*p* = 0.091	0.208	*p* = 0.287	-0.215	*p* = 0.254
	seed	0.320	*p* = 0.084	-0.331	*p* = 0.085	-0.288	*p* = 0.123
3-hydroxypropyl (3C)	leaves	NA	NA	NA	NA	NA	NA
	seed	-0.360	*p* = 0.050	-0.116	*p* = 0.555	0.360	*p* = 0.051
3-methylthiobutyl (3C)	leaves	NA	NA	NA	NA	NA	NA
	seed	-0.034	*p* = 0.857	-0.005	*p* = 0.981	0.016	*p* = 0.933
4-methylthiobutyl (4C)	leaves	NA	NA	NA	NA	NA	NA
	seed	**0.394**	***p* = 0.031**	0.025	*p* = 0.899	**-0.430**	***p* = 0.018**
7-methylsulfinylheptyl	leaves	NA	NA	NA	NA	NA	NA
	seed	**0.478**	***p* = 0.008**	0.177	*p* = 0.365	**-0.566**	***p* = 0.001**
7-methylthioheptyl	leaves	**0.423**	***p* = 0.019**	-0.012	*p* = 0.953	-0.197	*p* = 0.296
	seed	**0.570**	***p* = 0.001**	-0.156	*p* = 0.426	**-0.557**	***p* = 0.001**
Total 7C aliphatic glucosinolates	leaves	**0.487**	***p* = 0.006**	0.002	*p* = 0.991	-0.278	*p* = 0.137
	seed	**0.578**	***p* = 0.0008**	-0.090	*p* = 0.646	**-0.586**	***p* = 0.0006**
8-methylsulfinyloctyl	leaves	**0.474**	***p* = 0.008**	-0.078	*p* = 0.694	**-0.380**	***p* = 0.039**
	seed	0.178	*p* = 0.348	0.141	*p* = 0.473	-0.292	*p* = 0.117
8-methylthiooctyl	leaves	0.210	*p* = 0.266	-0.038	*p* = 0.846	0.012	*p* = 0.949
	seed	0.069	*p* = 0.719	-0.088	*p* = 0.655	-0.117	*p* = 0.539
Total 8C aliphatic glucosinolates	leaves	0.416	*p* = 0.022	-0.034	*p* = 0.862	-0.202	*p* = 0.284
	seed	0.115	*p* = 0.544	-0.007	*p* = 0.970	-0.196	*p* = 0.298
Total methylsulfinyl glucosinolates	leaves	0.469	*p* = 0.009	-0.017	*p* = 0.930	**-0.469**	***p* = 0.009**
	seed	0.232	*p* = 0.217	0.147	*p* = 0.454	-0.353	*p* = 0.056
Total aliphatic glucosinolates	leaves	0.159	*p* = 0.401	-0.027	*p* = 0.892	-0.144	*p* = 0.447
	seed	0.184	*p* = 0.330	-0.050	*p* = 0.801	-0.333	*p* = 0.072
Total glucosinolates	leaves	0.298	*p* = 0.109	-0.116	*p* = 0.556	-0.186	*p* = 0.326
	seed	0.162	*p* = 0.391	-0.061	*p* = 0.757	-0.341	*p* = 0.065

All three measures of RGR were tested for association with herbivore damage and with the line-specific glucosinolate concentrations in both leaves and seeds using Pearson's product moment correlation. Seed glucosinolate concentrations were also used because they might better reflect the lifetime production of glucosinolates by the plant and because the period of seed production is included in the growth curve. In contrast, leaf concentrations vary according to plant age and size and this can confound analyses (Koricheva, 1999). Concentrations of compounds were tested for normality and transformed where necessary (log or square-root) before correlations were performed.

Results

Conventional RGR

RGR_{early} was positively correlated with RGR_{av} (r = 0.389, P = 0.037, df = 27). As expected, there was a significant negative association between RGR_{av} and seed mass ($F_{1,28}$ = 6.47, P = 0.017) although RGR_{early} was not significantly associated with seed mass ($F_{1,27}$ = 1.99, P = 0.17).

Size-standardised RGR

The asymptotic regression model appeared to provide a good fit to the data (Fig. S1) and model-checking plots revealed no obvious signs of model mis-specification. As judged by comparison of AIC values (a measure of goodness-of-fit; Akaike (1974)), models with pot volume fitted as a factor were better than those in which the

relationship between pot volume and parameters was assumed to be linear or log-linear (although parameters always increased or decreased systematically with pot size). In larger pots the estimated asymptotic mass, A_i was higher ($F_{2,1039}$ = 9.23, p = 0.0001, Table 2) but the rate parameter, r_i was slightly lower ($F_{2,1039}$ = 16.02, p < 0.0001, Table 2). There was a significant effect of sown seed mass on the estimated initial mass, $M_{i,0}$, ($F_{1,987}$ = 17.18, p < 0.0001) which is expected if larger seeds produce larger seedlings. For the random effects, lines varied significantly only in the rate parameter, r_i (χ^2 = 20.9, df = 3, p < 0.0001); the asymptotic mass did not vary among lines (χ^2 = 3.16, df = 3, p = 0.368), nor was there any significant residual variation among lines in the estimated initial mass, $M_{i,0}$ once sown seed mass was fitted (χ^2 = 3.11, df = 1, p = 0.078). The lack of a genotype effect on asymptotic mass probably reflects the pot-grown conditions, in which final size is strongly limited by pot size (Paul-Victor & Turnbull, 2009). Size-standardised RGR was calculated for each line in the largest pot size using parameters taken from the final model and a reference mass, M_{ref} equal to the average mass of the plants half-way through the experiment (eqn 4). However, because only one parameter, r_i varied among lines, the relative ranking of lines with respect to growth rates is independent of the choice of reference mass. The relative ranking of genotypes is also independent of pot size, as there was no RIL × pot size interaction. In addition, this indicates that across-experiment comparisons are unlikely to be influenced by pot size differences. Size-standardised RGR was negatively correlated with RGR_{av} (r = -0.788, P < 0.0001, df = 29) and uncorrelated with RGR_{early} (r = -0.062, P = 0.749, df = 27).

Understanding relationships between seed size and RGR

To understand the relationships between seed size, conventional RGR and size-standardised RGR we show some simple results for the expected relationship between seed size and RGR assuming plant growth can be adequately modelled by the asymptotic regression equation above (eqn 2). In this case, we first assume that lines differ only in their seed mass ($M_{i,0}$) and that there are no true differences among lines in the two growth parameters (the rate parameter, r_i and the asymptotic mass, A_i). We thus assume that each line has an initial mass ($M_{i,0}$) given by its seed size and hence we can calculate the expected mass of each line at harvest 1 and harvest 6 using eqn 4 and average values of r_i and A_i estimated for the largest pot size (Table 2). We can then use these values to calculate RGR_{av} for each line (eqn.1). This reveals that while RGR_{av} is negatively correlated with seed size, size-standardised RGR (eqn 4) is the same for all lines (Fig. 2a and 2b). As other growth parameters are identical among lines, the variation in conventional RGR is entirely due to differences in initial

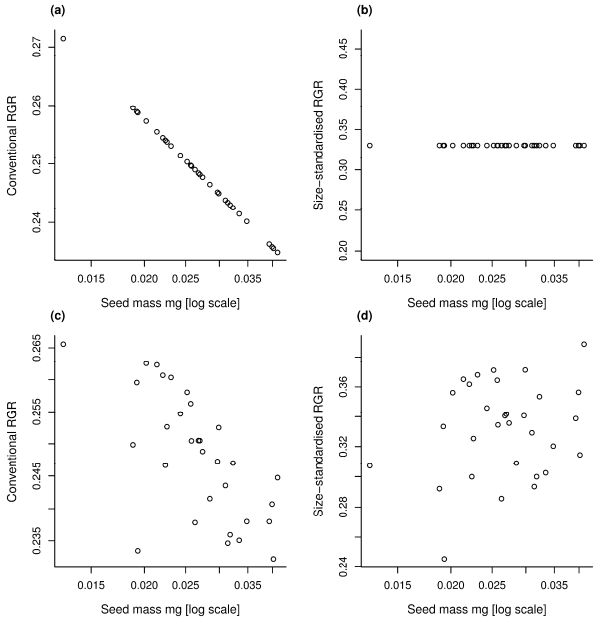


Figure 2 Expected relationships between seed size and RGR assuming plant growth can be adequately modelled by an asymptotic regression equation. In a and b we assume that there are no differences in the parameters of the growth rate equation other than differences in initial mass, while in c and d we incorporate the estimated among-line differences in the rate parameter, r_i.

mass, demonstrating that conventional RGR is sensitive to these differences. In contrast, size-standardised RGR correctly identifies that the growth parameters are identical.

Secondly, we can see the effect of including the line-specific differences in the rate parameter, r_i estimated by the model-fitting process. If we include these differences, the negative relationship between seed mass and conventional RGR persists (Fig. 2c), because RGR is very sensitive to differences in seed mass but relatively insensitive to differences in the growth parameter, r_i. Conventional RGR and size-standardised RGR are negatively correlated with each other because conventional RGR is negatively correlated with seed mass but there is a positive correlation between seed mass and size-standardised RGR. Thus lines with heavy seeds have low conventional RGR and high size-standardised RGR, while those with lighter seeds have high conventional RGR and low size-standardised RGR (Fig. 2c vs. 2d).

Secondary metabolites and herbivory

Among lines, damage by the specialist herbivore *P. xylostella* was positively correlated with size-standardised RGR, meaning that fast-growing lines suffered the most damage (Fig. 3a). In contrast, RGR_{av} was negatively correlated with damage suffered (fast-growing lines suffered the least damage; Fig. 3b) and RGR_{early} showed no correlation with herbivore damage (Fig. 3c). Correlations of damage by *T. ni* with RGR had the same direction as for *P. xylostella*, but were non-significant (not shown). Ten single glucosinolate compounds were assayed in the leaves,

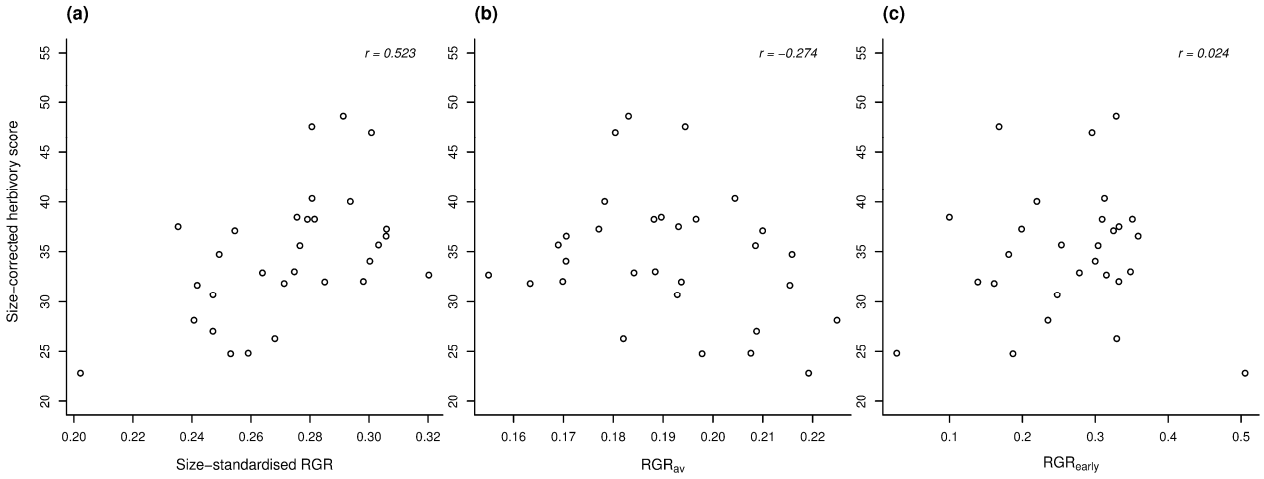


Figure 3 Residual herbivory means for *P. xylostella* plotted against the three measures of RGR. Values of r_P represent Pearson's product moment correlation. Only the correlation between herbivore damage and size-standardised RGR is significant.

the seeds or both (Table 3). As expected if defensive compounds have a growth cost, correlations between size-standardised RGR and ten glucosinolate compounds were predominantly negative in sign (Table 3). In fact, of the 26 correlations carried out in total, 23 were negative in sign. In contrast, correlations of chemical compounds with RGR_{av} were positive for all but one compound, while no correlation between chemical compounds and RGR_{early} was found (Table 3). Damage by the specialist herbivore was also negatively correlated with most glucosinolate compounds. Again out of 26 possible correlations, 23 were negative in sign although individual correlations were only significant in the case of the indolic glucosinolates (Table S2).

In general, correlations between size-standardised RGR and glucosinolate concentrations were stronger for seeds than for leaves, although not all compounds were measured in leaves, reducing the potential for significant correlations.

Discussion

The importance of size standardisation

RGR has for many years been accepted as a standardised way of measuring and comparing the growth rates of different species or genotypes. This is despite the fact that several authors have highlighted the problems with such comparisons when the species or genotypes vary in size (Poorter & Remkes, 1990; Reich *et al.*, 1998). Instantaneous RGR is expected to decline with size for both physiological reasons (large plants generally have to allocate more carbon to non-photosynthetic support tissue (Enquist *et al.*, 1999; West *et al.*, 2001)) and for reasons of resource restriction (large plants are increasingly unable to extract sufficient resources to maintain former growth rates). However, because the seed size variation observed in *Arabidopsis thaliana* is only 2–3 fold, it might reasonably be asked whether this variation is large enough to cause a problem. Here we have demonstrated that the seed mass variation in the L*er* x Cvi population is sufficient to generate a spurious negative correlation between conventional RGR and seed size, even when there is no true underlying variation in physiological growth rates. Thus, the method could be more widely used to disentangle the effects of size from other physiological differences among lines, not just those differences associated with defence (Coleman & McConnaughay, 1995; McConnaughay & Coleman, 1999; Bernacchi *et al.*, 2000). It should also be noted that non-destructive methods of measuring plant biomass (or leaf area) are increasingly available, perhaps removing the need for destructive harvests and hence avoiding some of the additional work associated with this method (e.g. see Durham Brooks *et al.*, 2009 for a new method of measuring root growth continuously).

The costs of defence

Traditionally, the costs of enhanced investment in defence have been assessed by comparisons of final seed set, as this is more directly correlated with fitness (Bazzaz *et al.*, 1987; Purrington & Bergelson, 1997; Mauricio, 1998). However, in the *Arabidopsis* lines analysed here, genotypes differed in the rate at which the asymptote was approached, and not in the asymptotic mass. Hence, it could be argued that a reduction in early growth rate does not represent a true fitness cost. However, a reduction in early growth rate could translate into a substantial fitness cost when plants are growing in competition rather than alone in individual pots. Rapid early growth allows resource pre-emption and therefore might be a good surrogate for competitive ability in short-lived annual plants (Grime, 2002). Loss of competitive status as a result of allocation to defence instead of early growth is a mechanism sometimes described as an 'opportunity cost' (Coley *et al.*, 1985), which is more easily detected when plants are growing in competition (reviewed in Koricheva, 2002).

Using the new RGR methodology, we were able to show that size-standardised growth rates were negatively correlated with a variety of glucosinolate compounds but positively correlated with herbivore damage. This supports the basic assumptions of plant defence theory which assumes that the optimal level of defence represents a balance between the costs of defence and the likelihood and severity of the expected attack (Herms & Mattson, 1992). In contrast, the direction of these correlations is reversed when using conventional RGR. Thus, if conventional RGR is to be believed, we would conclude that faster-growing lines produce more secondary metabolites and suffer less damage from herbivores, in common with some other studies using conventional RGR (Almeida-Cortez *et al.*, 1999; Almeida-Cortez & Shipley, 2002). It thus seems that the lack of negative correlations between growth rates and defensive compounds in some published studies could at least partly be due to the lack of size-standardisation when calculating growth rates.

The negative correlations between size-standardised growth rate and glucosinolate concentration were stronger for some individual compounds than for total glucosinolates (Table 3). Similarly, glucosinolate compounds cannot be treated as a single defence mechanism as there is structural specificity to their effectiveness against various insects, as demonstrated in both lab and field studies (Bidart-Bouzat & Kliebenstein, 2008; Hansen *et al.*, 2008). However, glucosinolates are products of complex metabolic pathways (e.g. Halkier & Gershenzon, 2006) and it might be that cellular processes independent of defence and growth influence their concentration and thus mask the trade-off pattern. Correlations between growth rates and glucosinolate concentration for all compounds were usually stronger in the seeds than in the leaves. The concentration in the seeds might better reflect the life-time metabolic potential of the maternal plant as of all plant organs, seeds have the highest proportion of glucosinolates per unit dry weight; as such they might incur a particularly high cost to the maternal plant (Brown *et al.*, 2003). Additionally some seed glucosinolates are derived from leaf-glucosinolates (Kliebenstein *et al.*, 2007; Nour-Eldin & Halkier, 2009). As a consequence, seed glucosinolates may correlate better with the growth rate calculated here, as this measure of growth is calculated over the whole lifespan of the plants, including the period of seed production.

Methodological considerations

The use of a homozygous RIL population, in which genetic variation is stable, allowed us to combine data from the present experiment with data on herbivory and glucosinolates collected in other experiments on the same RIL population. This is a well-established concept that has allowed for cross-comparison across numerous experimental conditions for a given RIL population, e.g. identifying causal links between transcriptome and metabolome variation even though the experiments were separated by several years (West *et al.*, 2006; Keurentjes *et al.*, 2007; Sønderby *et al.*, 2007; Wentzell *et al.*, 2007; Hansen *et al.*, 2008; Keurentjes *et al.*, 2008; Fu *et al.*, 2009; Sulpice *et al.*, 2009). It has also been established that the main QTLs controlling glucosinolate structure and concentration within this population do not show extensive genotype × environment interactions within the rosette (Kliebenstein *et al.*, 2001a; Kliebenstein *et al.*, 2002). Thus, although the measurements were made at three different points in time, this simply decreases our statistical power to find significant effects and should not introduce potential bias. The strength of the negative trade-off between growth and defence could therefore be under-estimated.

Correlative analyses such as those presented in this paper are not causal and are most effective for generating new hypotheses. Our analysis also raises the possibility that some of the costs of secondary metabolites are masked by cellular processes not directly associated with growth and defence. In *Arabidopsis*, such hypotheses require more rigorous testing with additional, larger RIL populations or by exploiting mutational variation. We hope, however, that the methods presented in the paper will better allow future studies to better estimate the costs of defence.

Acknowledgements

We would like to thank Theres Zwimpfer and Alex Schwendener for help in the glasshouse during the growth experiment and Nancy Bunbury and Carolina Huelva-Osuna for help with harvesting and weighing. This work was funded by Swiss National Science Foundation Grant 31-107531 to LAT and by National Science Foundation Grants DBI0642481 and MCB0323759 to DJK. TZ was partially supported by the Forschungskredit of the University of Zürich (grant 57230502).

References

Agrawal AA, Conner JK, Stinchcombe JR. 2004. Evolution of plant resistance and tolerance to frost damage. *Ecology Letters* **7**(12): 1199-1208.

Akaike H. 1974. A new look at the statistical model identification. *Ieee Transactions on Automatic Control* **AC19**(6): 716-723.

Almeida-Cortez JS, Shipley B, Arnason JT. 1999. Do plant species with high relative growth rates have poorer chemical defences? *Functional Ecology* **13**(6): 819-827.

Almeida-Cortez JS, Shipley W. 2002. No significant relationship exists between seedling relative growth rate under nutrient limitation and potential tissue toxicity. *Functional Ecology* **16**(1): 122-127.

Alonso-Blanco C, Blankestijn-de Vries H, Hanhart CJ, Koornneef M. 1999. Natural allelic variation at seed size loci in relation to other life history traits of

Arabidopsis thaliana. *Proc Natl Acad Sci U S A* **96**(8): 4710-4717.

Alonso-Blanco C, Peeters AJ, Koornneef M, Lister C, Dean C, van den Bosch N, Pot J, Kuiper MT. 1998. Development of an AFLP based linkage map of Ler, Col and Cvi Arabidopsis thaliana ecotypes and construction of a Ler/Cvi recombinant inbred line population. *Plant J* **14**(2): 259-271.

Arendt JD. 1997. Adaptive intrinsic growth rates: An integration across taxa. *Quarterly Review of Biology* **72**(2): 149-177.

Arendt JD. 2000. Allocation of cells to proliferation vs. differentiation and its consequences for growth and development. *Journal of Experimental Zoology* **288**(3): 219-234.

Barth C, Jander G. 2006. Arabidopsis myrosinases TGG1 and TGG2 have redundant function in glucosinolate breakdown and insect defense. *Plant Journal* **46**(4): 549-562.

Bazzaz FA, Chiariello NR, Coley PD, Pitelka LF. 1987. Allocating resources to reproduction and defense. *Bioscience* **37**(1): 58-67.

Beekwilder J, van Leeuwen W, van Dam NM, Bertossi M, Grandi V, Mizzi L, Soloviev M, Szabados L, Molthoff JW, Schipper B, et al. 2008. The impact of the absence of aliphatic glucosinolates on insect herbivory in Arabidopsis. *PLoS ONE* **3**(4): e2068.

Bergelson J, Purrington CB. 1996. Surveying patterns in the cost of resistance in plants. *American Naturalist* **148**(3): 536-558.

Bernacchi CJ, Coleman JS, Bazzaz FA, McConnaughay KDM. 2000. Biomass allocation in old-field annual species grown in elevated CO2 environments: no evidence for optimal partitioning. *Global Change Biology* **6**(7): 855-863.

Bidart-Bouzat MG, Kliebenstein DJ. 2008. Differential levels of insect herbivory in the field associated with genotypic variation in glucosinolates in *Arabidopsis thaliana*. *Journal of Chemical Ecology* **34**(8): 1026-1037.

Bones AM, Rossiter JT. 1996. The myrosinase-glucosinolate system, its organisation and biochemistry. *Physiologia Plantarum* **97**(1): 194-208.

Brown PD, Tokuhisa JG, Reichelt M, Gershenzon J. 2003. Variation of glucosinolate accumulation among different organs and developmental stages of *Arabidopsis thaliana*. *Phytochemistry* **62**(3): 471-481.

Coleman JS, McConnaughay KDM. 1995. A non-functional interpretation of a classical optimal-partitioning example. *Functional Ecology* **9**(6): 951-954.

Coley PD, Bryant JP, Chapin FS. 1985. Resource availability and plant antiherbivore defense. *Science* **230**(4728): 895-899.

Durham Brooks TL, Miller ND, Spalding EP. 2009. Plasticity of Arabidopsis root gravitropism throughout a multi-dimensional condition space quantified by automated image analysis. *Plant Physiol.*: pp.109.145292.

Enquist BJ, West GB, Charnov EL, Brown JH. 1999. Allometric scaling of production and life-history variation in vascular plants. *Nature* **401**(6756): 907-911.

Fu J, Keurentjes JJB, Bouwmeester H, America T, Verstappen FWA, Ward JL, Beale MH, de Vos RCH, Dijkstra M, Scheltema RA, et al. 2009. System-wide molecular evidence for phenotypic buffering in *Arabidopsis*. *Nature Genetics* **41**(2): 166-167.

Grime JP. 2002. *Plant Strategies, Vegetation Processes, and Ecosystem Properties*. Chichester: John Wiley & Sons, Ltd.

Halkier BA, Gershenzon J. 2006. Biology and biochemistry of glucosinolates. *Annual Review of Plant Biology* **57**: 303-333.

Hansen BG, Kerwin RE, Ober JA, Lambrix VM, Mitchell-Olds T, Gershenzon J, Halkier BA, Kliebenstein DJ. 2008. A novel 2-oxoacid-dependent dioxygenase involved in the formation of the goiterogenic 2-hydroxybut-3-enyl glucosinolate and generalist insect resistance in Arabidopsis. *Plant Physiology* **148**(4): 2096-2108.

Herms DA, Mattson WJ. 1992. The Dilemma of Plants - to Grow or Defend. *Quarterly Review of Biology* **67**(3): 283-335.

Hunt R. 1982. *Plant growth curves: a functional approach to plant growth analysis*. London: Edward Arnold.

Hunt R, Cornelissen JHC. 1997. Physiology, allocation, and growth rate: A reexamination of the Tilman model. *American Naturalist* **150**(1): 122-130.

Iwasa Y. 2000. Dynamic optimization of plant growth. *Evolutionary Ecology Research* **2**(4): 437-455.

Jander G, Cui JP, Nhan B, Pierce NE, Ausubel FM. 2001. The TASTY locus on chromosome 1 of Arabidopsis affects feeding of the insect herbivore *Trichoplusia ni*. *Plant Physiology* **126**(2): 890-898.

Keurentjes JJB, Fu JY, Terpstra IR, Garcia JM, van den Ackerveken G, Snoek LB, Peeters AJM, Vreugdenhil D, Koornneef M, Jansen RC. 2007. Regulatory network construction in *Arabidopsis* by using genome-wide gene expression quantitative trait loci. *Proceedings of the National Academy of Sciences of the United States of America* **104**(5): 1708-1713.

Keurentjes JJB, Sulpice R, Gibon Y, Steinhauser MC, Fu JY, Koornneef M, Stitt M, Vreugdenhil D. 2008. Integrative analyses of genetic variation in enzyme activities of primary carbohydrate metabolism reveal distinct modes of regulation in *Arabidopsis thaliana*. *Genome Biology* **9**(8): 20.

Kim JH, Jander G. 2007. *Myzus persicae* (green peach aphid) feeding on Arabidopsis induces the formation of a deterrent indole glucosinolate. *Plant Journal* **49**(6): 1008-1019.

Kliebenstein DJ. 2004. Secondary metabolites and plant/environment interactions: a view through

Arabidopsis thaliana tinged glasses. *Plant Cell and Environment* **27**(6): 675-684.

Kliebenstein DJ, D'Auria JC, Behere AS, Kim JH, Gunderson KL, Breen JN, Lee G, Gershenzon J, Last RL, Jander G. 2007. Characterization of seed-specific benzoyloxyglucosinolate mutations in *Arabidopsis thaliana*. *Plant Journal* **51**(6): 1062-1076.

Kliebenstein DJ, Gershenzon J, Mitchell-Olds T. 2001a. Comparative quantitative trait loci mapping of aliphatic, indolic and benzylic glucosinolate production in *Arabidopsis thaliana* leaves and seeds. *Genetics* **159**(1): 359-370.

Kliebenstein DJ, Kroymann J, Brown P, Figuth A, Pedersen D, Gershenzon J, Mitchell-Olds T. 2001b. Genetic control of natural variation in *Arabidopsis* glucosinolate accumulation. *Plant Physiology* **126**(2): 811-825.

Kliebenstein DJ, Pedersen D, Barker B, Mitchell-Olds T. 2002. Comparative analysis of quantitative trait loci controlling glucosinolates, myrosinase and insect resistance in *Arabidopsis thaliana*. *Genetics* **161**(1): 325-332.

Koricheva J. 1999. Interpreting phenotypic variation in plant allelochemistry: problems with the use of concentrations. *Oecologia* **119**(4): 467-473.

Koricheva J. 2002. Meta-analysis of sources of variation in fitness costs of plant antiherbivore defenses. *Ecology* **83**(1): 176-190.

Mauricio R. 1998. Costs of resistance to natural enemies in field populations of the annual plant *Arabidopsis thaliana*. *American Naturalist* **151**(1): 20-28.

McConnaughay KDM, Coleman JS. 1999. Biomass allocation in plants: Ontogeny or optimality? A test along three resource gradients. *Ecology* **80**(8): 2581-2593.

Nour-Eldin HH, Halkier BA. 2009. Piecing together the transport pathway of aliphatic glucosinolates. *Phytochemistry Reviews* **8**(1): 53-67.

Paul-Victor C, Turnbull LA. 2009. The effect of growth conditions on the seed size/number trade-off. *PLoS ONE* **4(9): e6917**: doi:10.1371/journal.pone.0006917.

Perrin N, Sibly RM. 1993. Dynamic-Models of Energy Allocation and Investment. *Annual Review of Ecology and Systematics* **24**: 379-410.

Pinheiro JC, Bates DM. 2000. *Mixed-Effects Models in S and S-Plus*. London: Springer.

Poorter H, Remkes C. 1990. Leaf-Area Ratio and Net Assimilation Rate of 24 Wild-Species Differing in Relative Growth-Rate. *Oecologia* **83**(4): 553-559.

Poorter L, Kitajima K. 2007. Carbohydrate storage and light requirements of tropical moist and dry forest tree species. *Ecology* **88**(4): 1000-1011.

Purrington CB, Bergelson J. 1997. Fitness consequences of genetically engineered herbicide and antibiotic resistance in Arabidopsis thaliana. *Genetics* **145**(3): 807-814.

R Development Core Team 2008. R: A Language and Environment for Statistical Computing.In. Vienna, Austria: R Foundation for Statistical Computing.

Reich PB, Tjoelker MG, Walters MB, Vanderklein DW, Bushena C. 1998. Close association of RGR, leaf and root morphology, seed mass and shade tolerance in seedlings of nine boreal tree species grown in high and low light. *Functional Ecology* **12**(3): 327-338.

Reichelt M, Brown PD, Schneider B, Oldham NJ, Stauber E, Tokuhisa J, Kliebenstein DJ, Mitchell-Olds T, Gershenzon J. 2002. Benzoic acid glucosinolate esters and other glucosinolates from *Arabidopsis thaliana*. *Phytochemistry* **59**(6): 663-671.

Rose KE, Atkinson RL, Turnbull LA, Rees M. 2009. The costs and benefits of fast living. *Ecology Letters* **12**: 1379-1384.

Siemens DH, Garner SH, Mitchell-Olds T, Callaway RM. 2002. Cost of defense in the context of plant competition: *Brassica rapa* may grow and defend. *Ecology* **83**(2): 505-517.

Sønderby IE, Hansen BG, Bjarnholt N, Ticconi C, Halkier BA, Kliebenstein DJ. 2007. A systems biology approach identifies a R2R3 MYB gene subfamily with distinct and overlapping functions in regulation of aliphatic glucosinolates. *Plos One* **2**(12): 16.

Sulpice R, Pyl ET, Ishihara H, Trenkamp S, Steinfath M, Witucka-Wall H, Gibon Y, Usadel B, Poree F, Piques MC, et al. 2009. Starch as a major integrator in the regulation of plant growth. *Proceedings of the National Academy of Sciences of the United States of America* **106**(25): 10348-10353.

Turnbull LA, Paul-Victor C, Schmid B, Purves DW. 2008. Growth rates, seed size, and physiology: Do small-seeded species really grow faster? *Ecology* **89**(5): 1352-1363.

Wentzell AM, Rowe HC, Hansen BG, Ticconi C, Halkier BA, Kliebenstein DJ. 2007. Linking metabolic QTLs with network and cis-eQTLs controlling biosynthetic pathways. *Plos Genetics* **3**(9): 1687-1701.

West GB, Brown JH, Enquist BJ. 2001. A general model for ontogenetic growth. *Nature* **413**(6856): 628-631.

West MAL, van Leeuwen H, Kozik A, Kliebenstein DJ, Doerge RW, St. Clair DA, Michelmore RW. 2006. High-density haplotyping with microarray-based expression and single feature polymorphism markers in *Arabidopsis*. *Genome Research* **16**(6): 787-795.

SUPPLEMENTARY INFORMATION

Table S1: Information about the 32 lines selected for the study. The two accessions L*er* and Cvi are the parents. The 30 remaining recombinant inbred lines are derived from reciprocal crosses between the two parents.

NASC	RIL Koornneef	Published Seed Mass (*) [mg]	Sown Seed mass (**) [mg]	*ERECTA* mutation
N8581	L*er*	0.0193	0.0202	1
N8580	Cvi	0.0351	0.0348	0
N22002	CVL3	0.0162	0.0129	1
N22014	CVL15	0.0145	0.0193	0
N22018	CVL19	0.0251	0.0263	1
N22026	CVL27	0.0275	0.0270	1
N22030	CVL31	0.0295	0.0334	0
N22033	CVL34	0.0236	0.0297	0
N22036	CVL37	0.0325	0.0399	0
N22037	CVL38	0.0150	0.0188	0
N22038	CVL39	0.0202	0.0258	0
N22043	CVL44	0.0242	0.0285	0
N22051	CVL53	0.0327	0.0310	1
N22057	CVL60	0.0286	0.0393	1
N22059	CVL62	0.0190	0.0224	0
N22094	CVL124	0.0274	0.0252	1
N22095	CVL125	0.0200	0.0214	0
N22098	CVL128	0.0273	0.0274	0
N22099	CVL129	0.0243	0.0268	0
N22105	CVL135	0.0327	0.0348	1
N22107	CVL137	0.0302	0.0314	0
N22109	CVL139	0.0217	0.0231	0
N22112	CVL142	0.0315	0.0318	1
N22124	CVL154	0.0317	0.0323	0
N22128	CVL158	0.0373	0.0411	1
N22130	CVL160	0.0361	0.0402	1
N22132	CVL162	0.0256	0.0221	1
N22138	CVL168	0.0334	0.0299	0
N22148	CVL178	0.0207	0.0226	1
N22149	CVL179	0.0223	0.0243	1
N22156	CVL187	0.0183	0.0192	1
N22160	CVL191	0.0280	0.0257	1

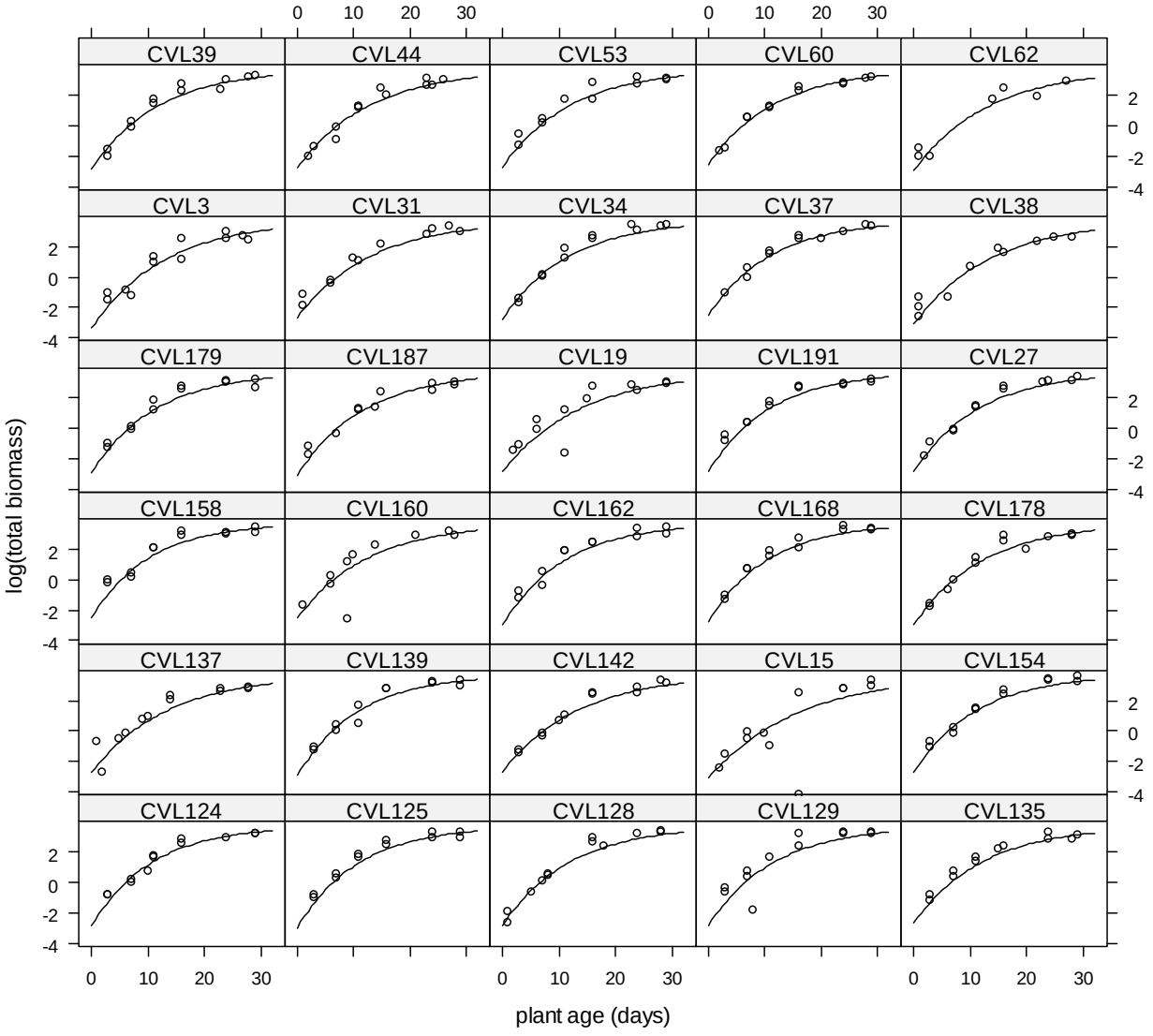
(*) Source: Alonso-Blanco et al., 1999.

(**) Source: Arabidopsis center (TAIR).

Table S2: Coefficients of Pearson's product moment correlation describing the relation of individual glucosinolates in leaves and seeds with herbivore damage of the specialist *Plutella xylostella* (size-corrected). Concentrations of compounds were log- or squareroot-transformed where necessary to meet the assumptions of normality.

		Herbivore damage *P. xylostella*	
indolyl-3-methyl	leaves	-0.216	$p = 0.253$
	seed	-0.323	$p = 0.082$
1-methoxy-indolyl-3-methyl	leaves	NA	NA
	seed	-0.130	$p = 0.493$
4-methoxy-indolyl-3-methyl	leaves	-0.135	$p = 0.479$
	seed	NA	NA
Total indolic glucosinolates	leaves	-0.144	$p = 0.447$
	seed	**-0.393**	$\mathbf{p = 0.032}$
3-hydroxypropyl (3C)	leaves	NA	NA
	seed	0.284	$p = 0.128$
3-methylthiobutyl (3C)	leaves	NA	NA
	seed	-0.103	$p = 0.590$
4-methylthiobutyl (4C)	leaves	NA	NA
	seed	-0.211	$p = 0.263$
7-methylsulfinylheptyl	leaves	NA	NA
	seed	-0.320	$p = 0.085$
7-methylthioheptyl	leaves	-0.081	$p = 0.670$
	seed	-0.306	$p = 0.100$
Total 7C aliphatic glucosinolates	leaves	-0.145	$p = 0.443$
	seed	-0.323	$p = 0.081$
8-methylsulfinyloctyl	leaves	-0.319	$p = 0.086$
	seed	-0.147	$p = 0.439$
8-methylthiooctyl	leaves	-0.048	$p = 0.800$
	seed	0.041	$p = 0.831$
Total 8C aliphatic glucosinolates	leaves	-0.179	$p = 0.343$
	seed	-0.031	$p = 0.873$
Total methylsulfinyl glucosinolates	leaves	-0.197	$p = 0.297$
	seed	-0.217	$p = 0.250$
Total aliphatic glucosinolates	leaves	-0.299	$p = 0.109$
	seed	-0.102	$p = 0.590$
Total glucosinolates	leaves	-0.215	$p = 0.254$
	seed	-0.143	$p = 0.451$

Figure S1. Destructive harvest data for the thirty RILs (the L*er* parent is not shown) grown in one pot size (pot diameter = 30 mm) with fitted curves from the final model.



Proc. R. Soc. B (2011) **278**
final accepted version

Using knockout mutants to reveal the growth costs of defensive traits

Tobias Züst[1,*], Bindu Joseph[3], Kentaro K. Shimizu[2], Daniel J. Kliebenstein[3] and Lindsay A. Turnbull[1]

[1]*Institute of Evolutionary Biology and Environmental Studies, University of Zürich, Winterthurerstrasse 190, Zürich 8057, Switzerland*
[2]*Institute of Plant Biology, Zollikerstrasse 107, University of Zürich, Zürich 8008, Switzerland*
[3]*Department of Plant Sciences, University of California at Davis, Davis, California 95616, USA*

We used a selection of *Arabidopsis thaliana* mutants with knockouts in defence genes to demonstrate growth costs of trichome development and glucosinolate production. Four of the seven defence mutants had significantly higher size-standardised growth rates (SGR) than the wildtype in early life, although this benefit declined as plants grew larger. SGR is known to be a good predictor of success under high-density conditions, and we confirmed that mutants with higher growth rates had a large advantage when grown in competition. Despite the lack of differences in flowering-time genes, the mutants differed in flowering time, a trait strongly correlated with early growth rate. Aphid herbivory decreased plant growth rate and increased flowering time, and aphid population growth rate was closely coupled to the growth rate of the host plant. Small differences in early SGR thus had cascading effects on both flowering time and herbivore populations.

Keywords: herbivore defence; size-standardised growth rate; glucosinolates; trichomes; *Arabidopsis*

1. INTRODUCTION

Plants deter herbivores through physical structures such as spines, thorns and hairs that reduce damage to leaf tissue [1, 2] and by producing toxic chemical compounds that reduce the growth rate or reproductive output of their enemies [3]. Such defences are assumed to be costly as they divert the plant's resources away from growth and reproduction [4-6]. However, experimental studies addressing fitness/defence trade-offs frequently fail to find the expected negative correlations [7-10], raising the question of whether such trade-offs are absent in many organisms (possibly through mechanisms which alleviate costs while maintaining resistance), or whether the methods employed to find them are inadequate [11].

Arabidopsis thaliana (L.) is attacked by a variety of pathogens [12] and herbivores, which include leaf-chewing caterpillars, sap-sucking aphids, flea beetles and leaf miners [13, 14]. As defence against these herbivores, *Arabidopsis* produces leaf hairs, called trichomes, and glucosinolates, a group of secondary metabolites [13]. Glucosinolate compounds are produced by all species of the Brassicaceae [15] and plants show large variation for this trait in the field [16], most likely as a consequence of differential selection by herbivore communities [17]. The majority of glucosinolates either have aliphatic or indolic side-chains [18]. Both types of glucosinolates negatively affect generalist leaf-chewing herbivores while aliphatic glucosinolates tend to affect these herbivores more severly [19-22]. Phloem-feeding aphids are mainly impaired by indolic glucosinolates [23] although there is evidence from field studies that some aphid species are also impaired by aliphatic glucosinolates [24]. Previously, we demonstrated that the production of glucosinolate compounds appeared to be costly to the plant, as there was a negative correlation between plant growth rate and glucosinolate content [11]. We also showed that slow-growing plants suffered reduced herbivore damage. While suggestive, these correlations are not proof of causal relationships. Instead, the costs of defensive traits can be more directly estimated using knockout mutants, in which defence genes are disabled artificially. Ideally, knockout mutants only differ from the wildtype in target genes, and if mutant phenotypes are not more extreme than the phenotypes of naturally-occurring variants, we believe that such mutants can be used to address ecological questions.

In this study we compared the growth rate of mutants reduced in specific defence mechanisms with the wildtype. We conducted a multiple-harvest experiment and calculated size-standardised relative growth rates (SGR), for a range of plant sizes [see also 11, 25]. A reduction in early growth rate is a likely consequence of diverting resources to defence; however, it is possible that for isolated plants growing with no competition there will be no measurable reduction in final seed output. This could occur because the resources diverted to defence compounds early in life can be later reclaimed and redirected to the seeds. However, under competitive conditions, a reduction in growth rate is likely to have severe fitness costs; for example, Fakheran et al. [26] showed that early growth rate was a very good predictor of success when a mixture of *Arabidopsis* genotypes were grown under high-density, competitive conditions. However, when grown alone, these same genotypes did not differ in their final biomass [11].

* Author for correspondence (tobias.zuest@cornell.edu)

In this study, we compared the growth rates of nine mutants with the wildtype in the presence and absence of the generalist aphid *Myzus persicae*. We also compared the growth rate of the aphid population on each of the ten genotypes and related this to the plant growth rate. Finally, we grew a subset of the genotypes in competition to test whether differences in early growth rates had greater fitness consequences under competitive conditions.

2. MATERIALS AND METHODS

(a) *Knockout mutants*

We used knockout mutants created in the genetic background of the *Arabidopsis* accession Columbia (Col-0, see Table S1 for a description of mutant phenotypes). One mutant (*gl1-2*) was originally created by x-ray mutagenesis and is deficient in trichome formation: the early leaves are entirely glabrous and there is greatly reduced trichome density on later leaves compared with the wildtype [27]. The *gl1-2* mutant also shows decreased phenolic defence expression (Daniel J. Kliebenstein, unpublished data). All other mutants were originally created by T-DNA insertion. The mutants *myb28*, *myb29* and *myb28myb29* contain knockouts in transcription factors that decrease expression of aliphatic glucosinolates [21, 28] and the genotypes *cyp79B2*, *cyp79B3* and *cyp79B2cyp79B3* contain enzyme knockouts that decrease or abolish the indolic glucosinolate and camalexin pathways [29]. The genes *MYB28*/*MYB29* and *CYP79B2*/*CYP79B3* are tandem duplicated genes within their respective cellular pathway and are traditionally considered redundant [28, 29]. To control for non-target effects of transgenic plants such as the cost of expression of selection marker genes, we included two genotypes with knockouts in genes not associated with defence and with no predicted fitness costs: *ppi1-2* and *nst1-2* [30, 31].

Even though all mutants used here were originally created by artificial gene knockout, similar phenotypes can be found in natural accessions of *Arabidopsis*. For example, the accessions est-0 (NASC 1148) and wil-3 (NASC 1598) are both completely glabrous, and glucosinolate levels vary considerably among natural accessions [32].

(b) *Experimental design*

Plants were grown in a mixture of peat- based substrate (PP7, Tref Group, The Netherlands) and sand in a ratio 1:1. Each pot (diameter = 40 mm, depth = 70 mm) was sown with five seeds and cold stratified at 4° C for 48 hours. The pots were then moved to a glasshouse with supplemental artificial light at a 16h light / 8h dark cycle and 26° C day / 22° C night temperature. Plants were watered twice a week throughout the experiment and no additional nutrients were supplied. Five days after sowing, seedlings were thinned to leave only the most central seedling. Bolting (initiation of the flowering stem) was recorded for each plant to the nearest day. Six plants per genotype were harvested on days 5, 9, 13, 18, 23, 29 and 35 after germination. On day 5, the herbivore treatment was initiated by placing a single 1st instar aphid onto half of the remaining plants. The offspring of the introduced aphids (F_1) were counted and removed at each harvest to keep herbivore pressure roughly constant among plant genotypes.

(c) *Size-standardised RGR of plants*

We fitted an asymptotic regression model log(aboveground biomass) through time:

$$\log(M_{i,t}) = A_i + (\log(M_{i,0}) - A_i)\exp(-\exp(r_i)t) \qquad (2.1)$$

where $M_{i,0}$ is the starting mass of genotype i at $t = 0$, A_i is the asymptotic mass as $t \to \infty$ and r_i is the logarithm of the rate constant. The model was fitted with the function *gnls* in R [33] with genotype treated as a fixed effect. Models were compared based on their AIC values and size-standardised growth rates (SGR) were calculated with parameters taken from the best model. SGR is given by

$$SGR_i = \exp(r_i)(A_i - \log(M_{ref})) \qquad (2.2)$$

where M_{ref} is a reference mass (for derivation of equation 2.2 see Appendix S1 and [25]).

(d) *Prediction intervals on SGR*

Gnls produces point estimates and confidence intervals for the two estimated model parameters, the rate constant r_i and the asymptotic mass A_i. To estimate confidence intervals for SGR (a function of these two parameters), we generated population prediction intervals [34, 35]. The method assumes that the distribution of the parameters is multivariate normal with a variance-covariance matrix given by the inverse of the information matrix. We used the function *mvrnom*, which selects multivariate normal random deviates, and the variance-covariance matrix given by the function *vcov*. We generated 1000 sets of parameters to calculate a distribution of differences between wildtype and mutant SGRs. The lower and upper 95% quantile of these distributions are the boundaries of the prediction intervals. Mutant SGRs are significantly different from wildtype SGR if the prediction interval does not include zero. Point estimates of SGR and prediction intervals were calculated at two reference masses (M_{ref}, equation 2.2): an early SGR using the average mass at age = 5 days and a late SGR using the average mass at age = 29 days.

(e) *Aphid rate of reproduction*

Aphid performance was analysed by fitting the same asymptotic model (equation 2.1) to the log-transformed cumulative number of F_1 aphids, thus generating a size-standardised relative growth rate of the aphid population. Estimates and prediction intervals of aphid SGR were calculated at two reference population sizes: 2 and 42 individuals, roughly corresponding to average offspring number on day 13 and 29.

(f) *Early growth rate and competition*

To determine whether differences in early growth rate affected the outcome of competition, we carried out a competition experiment with a subset of genotypes: *myb28*, *myb29* and the wildtype. Plants were grown in 5 x 5.5 cm pots filled with germination soil and maintained under long day (16h light / 8h dark) conditions in a controlled environment growth chamber. Prior to sowing, seeds were imbibed and cold stratified at 4° C for 3 days. In each pot, nine seeds were arranged into a square with an area of 1 cm^2, thus closely

surrounding the central seed with eight neighbours. Mutant central seeds were either surrounded by their own genotype. or by the wildtype, while wildtype central seeds were surrounded by *myb28*, *myb29* or wildtype, resulting in a total of seven combinations. Each combination was replicated 12 times, half of which were harvested after three weeks and half after four weeks. There was some germination failure and only pots with more than 5 neighbour plants were kept, thus the sample size was decreased to 31 pots in week 3 and 28 pots in week 4. At day 18 for week 3 and day 25 for week 4, the rosette diameter of the central plant and two neighbours was recorded. Three days later, the same plants were harvested and fresh weight was measured. Fresh weight or rosette diameter were analysed as a function of target genotype, neighbour genotype and harvest week using linear models.

3. RESULTS

(a) *SGR of plant genotypes*

The final asymptotic regression model included effects of plant genotype and herbivory on the rate constant r_i and the asymptotic mass A_i as judged by comparing AIC values (Table S2, Figure S1). There was no herbivory × plant genotype interaction. For the following analysis, only results from the control (without aphids) are shown.

Six of the seven defence mutants had significantly higher values of the rate constant r_i than wildtype, while the two mutants with knockouts in other genes did not differ from wildtype (Table 1). In contrast, all mutants had lower values of the asymptotic mass A_i compared to wildtype (Table S2). Early SGR was significantly higher than wildtype for the glabrous mutant *gl1-2*, the indole glucosinolate mutants *cyp79B3* and *cyp79B2cyp79B3* and the aliphatic glucosinolate mutant *myb28* (Figure 1a). In later life, mutants tended to have equal or lower SGRs than the wildtype (Figure 1b).

As an unexpected result, we found that across the ten genotypes, early SGR is an excellent predictor of mean bolting age (r = -0.813; $F_{1,8}$ = 15.63, p = 0.004), i.e., fast-growing genotypes flowered earlier. This demonstrates that changes in early growth rate can influence flowering time, despite the fact that the mutant genotypes in question did not contain altered flowering genes. This apparently direct link between early growth rate and flowering time is confirmed by the aphid treatment: aphid feeding also decreased growth rate but increased bolting age in all genotypes (Table 1).

(b) *Aphid rate of reproductive output*

The asymptotic regression model included effects of plant genotype on the rate constant r_i and the asymptotic mass A_i (Figure S2, Table S3). With the exceptions of *ppi1-2* and *nst1-2*, none of the aphid SGRs calculated from this model were significantly different from wildtype (Figure 1c, d). However, the aphid rate of reproductive output on the different plant genotypes was strongly correlated with plant SGR at early stages (r = 0.877, $F_{1,8}$ = 26.67, p = 0.0009), and this correlation, even though weakened, was still present at the end of the experiment (r = 0.630, $F_{1,8}$ = 5.26, p = 0.051). Thus, aphid populations performed better on fast-growing genotypes.

(c) *SGR and competition*

Based on measurements of early SGR, we would predict that *myb28* should outcompete the wildtype, whereas *myb29* and wildtype should be equal competitors. In the analysis of fresh weight, neighbour genotype had a significant effect on the target genotype ($F_{2,23}$ = 5.74, p = 0.010). In week 4, *myb28* target plants weighed 0.18 (± 0.06, 1SE) grams when surrounded by other *myb28* plants, but weighed 0.41 (± 0.07, 1SE) grams when surrounded by wildtype plants. Wildtype plants surrounded by wildtype neighbours weighed on average 0.29 (± 0.07, 1SE) grams, while wildtype plants surrounded by *myb28* neighbours weighed only 0.09 (± 0.07, 1SE) grams. The weight of *myb29* was not significantly affected by neighbour identity. The direction of the effects in week 3 and for rosette diameter in both weeks was similar but non-significant. Thus, it seems that the observed significant difference in early growth rate between *myb28* and wildtype has fitness consequences when the plants are grown in competition.

4. DISCUSSION

Six of the seven genotypes with knockouts in defence genes had a higher rate constant (r_i) than the wildtype but the asymptotic mass (A_i) was lower for all mutants. As SGR is a function of both parameters, this meant that only four defence mutants had significantly higher early growth rate than the wildtype, and this difference decreased with increasing plant size. The observed differences in early growth rate were relatively small, but these differences had large effects on target plant size when growing in competition. For example, *myb28* has a higher initial growth rate than wildtype and thus should be able to outcompete it when the two genotypes are grown together. In support of this, *myb28* was more than twice as large with wildtype as with *myb28* neighbours and similarly, wildtype individuals were larger with wildtype than with *myb28* neighbours. In contrast, the early growth rate of *myb29* (which was only grown with either wildtype or *myb29* neighbours in the competition experiment) is similar to wildtype and it was unaffected by neighbour identity when grown under competition. The large advantage observed under competitive conditions is not unexpected under scramble competition for resources, as a difference in early growth rate will lead to unequal resource uptake, and with a finite pool of resources, the plant with the higher uptake rate will gain a greater share of the total. In a recent study, Fakheran et al. [26] also showed that early growth rate was the best predictor of success in high-density competitive landscapes. Differences in growth rates among genotypes are thus also likely to be the underlying mechanism creating the sometimes ambiguous results from studies looking at kinship effects on competitive ability of plants [e.g. 36, 37].

Early growth rate was also a very good predictor of flowering time, a trait that varied by several days among genotypes, despite identical flowering genes. Aphid herbivory also reduced early growth rate and increased flowering time, again indicating a possible causal link between early growth rate and the decision to flower. Small differences in early growth rate are therefore biologically relevant, leading to a disadvantage in competition and to delayed flowering. Hence the production of defensive traits, and the consequent reduction in growth rate are likely to be costly to the plant.

Table 1. Parameters from the asymptotic regression model and bolting age of plant genotypes. Parameters for the wildtype are absolute values while the parameter values of mutants are differences from the wildtype. Bolting ages are absolute values. Aphid gives the overall difference in parameter values or age at bolting in the presence of aphids. Significant differences are in boldface.

plant genotype	*rate constant (r_i)*	*asymptotic mass (A_i)*	bolting age
Wildtype	-2.25	3.58	18.6
gl1-2	**+0.12**	-0.27	**17.7**
cyp79B2	-0.00	-0.26	**19.8**
cyp79B3	**+0.18**	**-0.30**	18.0
cyp79B2cyp79B3	**+0.16**	**-0.36**	**16.4**
myb28	**+0.13**	-0.29	17.8
myb29	**+0.10**	**-0.37**	**17.3**
myb28myb29	**+0.12**	**-0.64**	**20.1**
nst1-2	+0.06	**-0.71**	**19.6**
ppi1-2	-0.04	-0.21	**22.2**
Aphid	-0.06	-0.14	**+0.47**

This supports findings from field experiments which show that both trichomes and glucosinolates have a visible fitness cost if herbivores are eliminated [e.g. 13]. It also supports theoretical work that assumes such a trade-off between defence and fitness.

Surprisingly, genotypes with knockouts in the homologous gene pairs MYB28/MYB29 and CYP79B2/CYP79B3 had relatively large differences in their growth rate. *cyp79B2* grew more slowly than *cyp79B3* and the double mutant, and *myb28* grew faster than *myb29* and the double *myb28*/*myb29* mutant. *MYB28* and *MYB29* are not completely functionally redundant and there is evidence of an incoherent feed-forward loop involving these two genes that complicates our ability to place them in a linear pathway [38]. Likewise, *CYP79B2* and *CYP79B3* are not completely functionally redundant, with the genes having quantitative preferences to the camalexin versus indole glucosinolate pathways. How the fluxes are reshuffled in the single mutants is not currently understood and as such, the double *CYP79B2B3* is a cleaner background to directly interpret [39]. These data suggest that the genes *MYB28*/*MYB29* and *CYP79B2*/*CYP79B3* are involved in non-linear pathways that are not completely understood and will require further research to parse. This does suggest that single gene mutants in any background may be more complicated to interpret than is traditionally considered.

Defence mutants benefited from the lack of defensive traits in early life; but, as plants grew larger, this benefit apparently disappeared. In contrast, the two mutants with knockouts in other (non-defence-related) genes performed worse than the wildtype at all sizes – a phenomenon that was not observed previously; hence these mutants were thought to be neutral [30, 31]. The poor performance of the two non-defence-related mutants in our study may be due to the growing conditions: our plants were grown in small pots in a sand/soil mixture with no additional nutrients, and this could be a more stressful environment than that normally used for genetic work. That all mutants had poorer performance at larger sizes is possibly due to pleiotropic effects, as disabling a gene usually affects several functions. It could also be due to the expression of selection marker genes, which might have associated costs (although this would not explain the poor performance of *gl1-2*, which is not a transgenic).

According to optimal defence theory [40] plants should follow different defence strategies before and after bolting, hence the decline in mutant SGRs with respect to wildtype could also represent a change in the value of defensive traits. Prior to bolting, growth is mass dependent and removal of leaf tissue by herbivores should be particularly costly, thus plants should invest heavily in leaf defences. Mutant plants, unable to produce such defensive traits, then have additional resources available for growth. After bolting, the inflorescence becomes the most valuable plant organ. However, at least part of the defensive compounds in the inflorescence are relocated from rosette leaves [18]; wildtype plants might thus synthesise less glucosinolates *de-novo* during the post-flowering period, hence decreasing the relative advantage of knockout mutants.

All plant genotypes were similarly susceptible to aphid herbivory and aphid performance was not generally better on genotypes with knockouts in defence genes. However, if aphids remove a constant fraction of the plant's resources, we still expect faster-growing plants to support higher aphid population growth (see Hautier et al. [41] for a similar situation with a parasitic plant, *Rhinanthus alectorolophus*). This was indeed the case, as aphid population growth rate was strongly correlated with plant SGR. The relatively small differences in aphid population size on wildtype and mutant plants in our study is probably partly a result of keeping aphid densities low by constantly removing offspring. Low herbivore densities might in turn be unable to trigger a full defensive response by the plants; as part of the defence response of *Arabidopsis* is only induced by herbivore feeding [23, 42, 43]. That high concentrations of certain glucosinolate compounds can affect aphid feeding has been shown by Kim & Jander [23], who demonstrated that indolic, but not aliphatic glucosinolates deterred *M. persicae* when applied in artificial diets. However, Kim et al. [44], too, failed to show increased aphid reproduction on the *cyp79B2cyp79B3* double-knockout mutant and only demonstrated decreased reproduction on a mutant overexpressing indolic glucosinolates. The specific mechanism involved in plant defence against aphids thus remains unclear, while the relevance of glucosinolates in defence against leaf-chewing herbivores has been demonstrated repeatedly [19-22].

In summary, mutants with knockouts in defence genes generally grew faster at small sizes than the wildtype. This enhanced early growth rate gave them an advantage in competition and allowed them to flower earlier. Combined with earlier work demonstrating a negative correlation between glucosinolate concentrations and growth rates, this study supports the hypothesis that the defence traits of *Arabidopsis* are costly to the plant. While knockout mutants helped to reveal these costs, such mutants can exhibit growth disadvantages, particularly in later life, and especially when grown under nutrient-poor conditions, and hence should be used with caution.

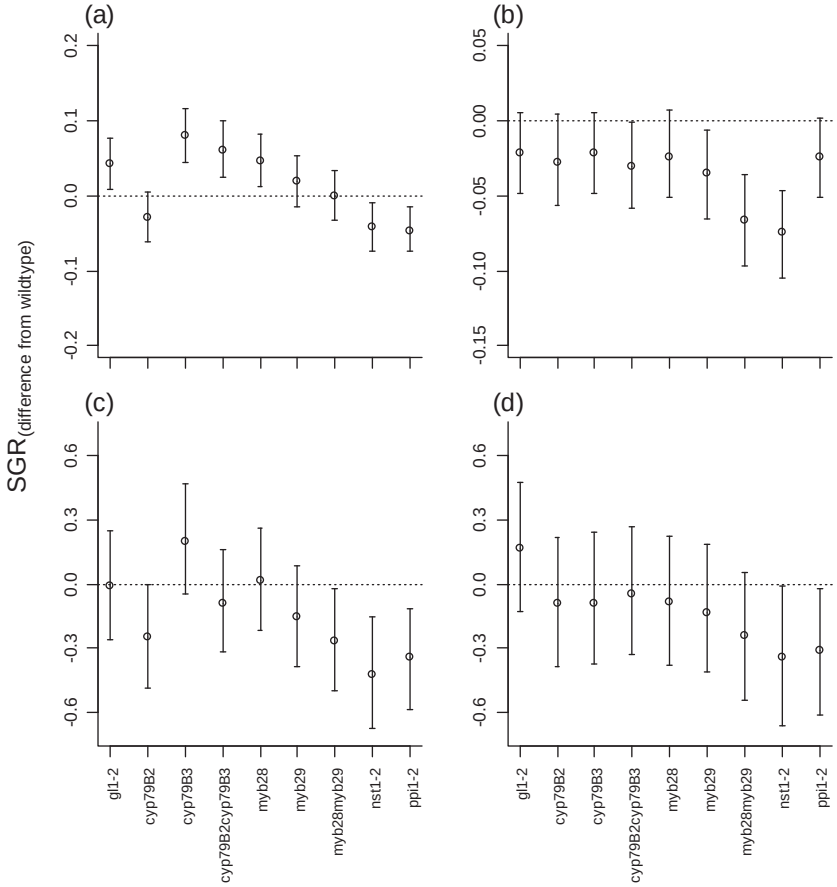


Figure 1. Differences in size-standardized relative growth rates (SGRs) of mutant plants from wild-type in (a,b) and (c,d) population SGRs of aphids feeding on mutant plants. For plants, early SGR is calculated for average mass (a) at age = 5 days and (b) at age = 29 days, while for aphids, SGR is calculated at the average population size (c) when plant age = 13 days and (d) when plant age = 29 days. Dotted lines represent zero difference from wild-type in SGR, error bars show 95% prediction intervals.

This work was funded by the Forschungskredit of the University of Zürich (grant 57230502) to TZ and by Swiss National Science Foundation Grant 31-107531 to LAT. Several colleagues kindly provided seeds (Table S1).

REFERENCES

1. Levin, D. A. 1973 The role of trichomes in plant defense. *Q. Rev. Biol.* **48,** 3-15.
2. Grubb, P. J. 1992 A positive distrust in simplicity - lessons from plant defences and from competition among plants and among animals. *J. Ecol.* **80,** 585-610.
3. Fraenkel, G. S. 1959 The raison d'être of secondary plant substances. *Science* **129,** 1466-1470.
4. Bazzaz, F. A., Chiariello, N. R., Coley, P. D. & Pitelka, L. F. 1987 Allocating resources to reproduction and defense. *Bioscience* **37,** 58-67.
5. Coley, P. D., Bryant, J. P. & Chapin, F. S. 1985 Resource availability and plant antiherbivore defense. *Science* **230,** 895-899.
6. Herms, D. A. & Mattson, W. J. 1992 The dilemma of plants: to grow or defend. *Q. Rev. Biol.* **67,** 283-335.
7. Koricheva, J. 2002 Meta-analysis of sources of variation in fitness costs of plant antiherbivore defenses. *Ecology* **83,** 176-190.
8. Bergelson, J. & Purrington, C. B. 1996 Surveying patterns in the cost of resistance in plants. *Am. Nat.* **148,** 536-558.
9. Almeida-Cortez, J. S., Shipley, B. & Arnason, J. T. 1999 Do plant species with high relative growth rates have poorer chemical defences? *Funct. Ecol.* **13,** 819-827.
10. Almeida-Cortez, J. S. & Shipley, W. 2002 No significant relationship exists between seedling relative growth rate under nutrient limitation and potential tissue toxicity. *Funct. Ecol.* **16,** 122-127.
11. Paul-Victor, C., Züst, T., Rees, M., Kliebenstein, D. J. & Turnbull, L. A. 2010 A new method for measuring relative growth rate can uncover the costs of defensive compounds in *Arabidopsis thaliana*. *New Phytol.* **187,** 1102-1111. (doi: 10.1111/j.1469-8137.2010.03325.x)
12. Tian, D., Traw, M. B., Chen, J. Q., Kreitman, M. & Bergelson, J. 2003 Fitness costs of R-gene-mediated resistance in *Arabidopsis thaliana*. **423,** 74-77.
13. Mauricio, R. 1998 Costs of resistance to natural enemies in field populations of the annual plant *Arabidopsis thaliana*. *Am. Nat.* **151,** 20-28.
14. Mauricio, R. & Rausher, M. D. 1997 Experimental manipulation of putative selective agents provides evidence for the role of natural enemies in the evolution of plant defense. *Evolution* **51,** 1435-1444.
15. Fahey, J. W., Zalcmann, A. T. & Talalay, P. 2001 The chemical diversity and distribution of glucosinolates and isothiocyanates among plants. *Phytochemistry* **56,** 5-51.
16. Mithen, R., Raybould, A. F. & Giamoustaris, A. 1995 Divergent selection for secondary metabolites between wild populations of *Brassica oleracea* and its

implications for plant-herbivore interactions. *Heredity* **75,** 472-484.
17. Newton, E. L., Bullock, J. M. & Hodgson, D. J. 2009 Glucosinolate polymorphism in wild cabbage (*Brassica oleracea*) influences the structure of herbivore communities. *Oecologia* **160,** 63-76. (doi: 10.1007/s00442-009-1281-5)
18. Brown, P. D., Tokuhisa, J. G., Reichelt, M. & Gershenzon, J. 2003 Variation of glucosinolate accumulation among different organs and developmental stages of *Arabidopsis thaliana*. *Phytochemistry* **62,** 471-481.
19. Burow, M., Müller, R., Gershenzon, J. & Wittstock, U. 2006 Altered glucosinolate hydrolysis in genetically engineered *Arabidopsis thaliana* and its influence on the larval development of *Spodoptera littoralis*. *J. Chem. Ecol.* **32,** 2333-2349. (doi: 10.1007/s10886-006-9149-1)
20. Lankau, R. A. 2007 Specialist and generalist herbivores exert opposing selection on a chemical defense. *New Phytol.* **175,** 176-184. (doi: 10.1111/j.1469-8137.2007.02090.x)
21. Beekwilder, J., van Leeuwen, W., van Dam, N. M., Bertossi, M., Grandi, V., Mizzi, L., Soloviev, M., Szabados, L., Molthoff, J. W., Schipper, B., et al. 2008 The impact of the absence of aliphatic glucosinolates on insect herbivory in Arabidopsis. **3,** e2068. (doi: 10.1371/journal.pone.0002068)
22. Schlaeppi, K., Bodenhausen, N., Buchala, A., Mauch, F. & Reymond, P. 2008 The glutathione-deficient mutant pad2-1 accumulates lower amounts of glucosinolates and is more susceptible to the insect herbivore *Spodoptera littoralis*. *Plant J.* **55,** 774-786. (doi: 10.1111/j.1365-313X.2008.03545.x)
23. Kim, J. H. & Jander, G. 2007 *Myzus persicae* (green peach aphid) feeding on Arabidopsis induces the formation of a deterrent indole glucosinolate. *Plant J.* **49,** 1008-1019. (doi: 10.1111/j.1365-313X.2006.03019.x)
24. Newton, E., Bullock, J. M. & Hodgson, D. 2009 Bottom-up effects of glucosinolate variation on aphid colony dynamics in wild cabbage populations. *Ecol. Entomol.* **34,** 614-623. (doi: 10.1111/j.1365-2311.2009.01111.x)
25. Rose, K. E., Atkinson, R. L., Turnbull, L. A. & Rees, M. 2009 The costs and benefits of fast living. *Ecol. Lett.* **12,** 1379-1384. (doi: 10.1111/j.1461-0248.2009.01394.x)
26. Fakheran, S., Paul-Victor, C., Heichinger, C., Schmid, B., Grossniklaus, U. & Turnbull, L. A. 2010 Adaptation and extinction in experimentally fragmented landscapes. *PNAS* **107,** 19120-19125. (doi: 10.1073/pnas.1010846107)
27. Esch, J. J., Oppenheimer, D. G. & Marks, M. D. 1994 Characterization of a weak allele of the GL1 gene of *Arabidopsis thaliana*. *Plant Mol. Biol.* **24,** 203-207.
28. Sønderby, I. E., Hansen, B. G., Bjarnholt, N., Ticconi, C., Halkier, B. A. & Kliebenstein, D. J. 2007 A systems biology approach identifies a R2R3 MYB gene subfamily with distinct and overlapping functions in regulation of aliphatic glucosinolates. *PLoS One* **2,** e1322. (doi: 10.1371/journal.pone.0001322)
29. Zhao, Y. D., Hull, A. K., Gupta, N. R., Goss, K. A., Alonso, J., Ecker, J. R., Normanly, J., Chory, J. & Celenza, J. L. 2002 Trp-dependent auxin biosynthesis in *Arabidopsis*: involvement of cytochrome P450s CYP79B2 and CYP79B3. *Genes Dev.* **16,** 3100-3112. (doi: 10.1101/gad.1035402)
30. Mitsuda, N., Iwase, A., Yamamoto, H., Yoshida, M., Seki, M., Shinozaki, K. & Ohme-Takagi, M. 2007 NAC transcription factors, NST1 and NST3, are key regulators of the formation of secondary walls in woody tissues of *Arabidopsis*. *Plant Cell* **19,** 270-280. (doi: 10.1105/tpc.106.047043)
31. Anzi, C., Pelucchi, P., Vazzola, V., Murgia, I., Gomarasca, S., Piccoli, M. B. & Morandini, P. 2008 The proton pump interactor (Ppi) gene family of *Arabidopsis thaliana*: expression pattern of Ppi1 and characterisation of knockout mutants for Ppi1 and 2. **10,** 237-249. (doi: 10.1111/j.1438-8677.2007.00022.x)
32. Kliebenstein, D. J., Kroymann, J., Brown, P., Figuth, A., Pedersen, D., Gershenzon, J. & Mitchell-Olds, T. 2001 Genetic control of natural variation in *Arabidopsis* glucosinolate accumulation. *Plant Physiol.* **126,** 811-825.
33. R Development Core Team 2010 *R: a language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing
34. Lande, R., Engen, S. & Sæther, B.-E. 2003 *Stochastic population dynamics in ecology and conservation*. Oxford, England: Oxford University Press.
35. Bolker, B. M. 2008 *Ecological models and data in R*. Princeton, New Jersey: Princeton University Press.
36. Tonsor, S. J. 1989 Relatedness and intraspecific competition in *Plantago lanceolata*. *Am. Nat.* **134,** 897-906.
37. Masclaux, F., Hammond, R. L., Meunier, J., Gouhier-Darimont, C., Keller, L. & Reymond, P. 2010 Competitive ability not kinship affects growth of *Arabidopsis thaliana* accessions. *New Phytol.* **185,** 322-331. (doi: 10.1111/j.1469-8137.2009.03057.x)
38. Sønderby, I. E., Burow, M., Rowe, H. C., Kliebenstein, D. J. & Halkier, B. A. 2010 A complex interplay of three R2R3 MYB transcription factors determines the profile of aliphatic glucosinolates in Arabidopsis[1[C][W][OA]]. *Plant Physiol.* **153,** 348-363. (doi: 10.1104/pp.109.149286)
39. Kliebenstein, D. J., Rowe, H. C. & Denby, K. J. 2005 Secondary metabolites influence Arabidopsis/*Botrytis* interactions: variation in host production and pathogen sensitivity. *Plant J.* **44,** 25-36. (doi: 10.1111/j.1365-313X.2005.02508.x)
40. Rhoades, D. F. 1979 Evolution of plant chemical defense against herbivores. In *Herbivores: their interaction with secondary metabolites* (ed. G. A. Rosenthal & D. H. Janzen), pp. 3-54. New York: Academic Press.
41. Hautier, Y., Hector, A., Vojtech, E., Purves, D. & Turnbull, L. A. 2010 Modelling the growth of parasitic plants. *J. Ecol.* **98,** 857-866. (doi: 10.1111/j.1365-2745.2010.01657.x)
42. Gigolashvili, T., Berger, B., Mock, H. P., Müller, C., Weisshaar, B. & Flügge, U. I. 2007 The transcription factor HIG1/MYB51 regulates indolic glucosinolate biosynthesis in *Arabidopsis thaliana*. *Plant J.* **50,** 886-901. (doi: 10.1111/j.1365-313X.2007.03099.x)
43. Mewis, I., Tokuhisa, J. G., Schultz, J. C., Appel, H. M., Ulrichs, C. & Gershenzon, J. 2006 Gene expression and glucosinolate accumulation in *Arabidopsis thaliana* in response to generalist and specialist herbivores of different feeding guilds and the role of defense signaling pathways. *Phytochemistry* **67,** 2450-2462. (doi: 10.1016/j.phytochem.2006.09.004)
44. Kim, J. H., Lee, B. W., Schroeder, F. C. & Jander, G. 2008 Identification of indole glucosinolate breakdown products with antifeedant effects on *Myzus persicae* (green peach aphid). *Plant J.* **54,** 1015-1026. (doi: 10.1111/j.1365-313X.2008.03476.x)

Appendix S1. Derivation of eqn 2.

Plant biomass is log-transformed giving:

$$\log\left(M_{i,t}\right) = A_i + (\log(M_{i,0}) - A_i)\exp(-\exp(r_i)t) \qquad \text{eqn S1}$$

The time required to reach a given reference mass, M_{ref}, is given by

$$t(M_{ref}) = \log\left(\frac{\log(M_{i,0}) - A_i}{\log(M_{ref}) - A_i}\right) \times \exp(-r_i) \qquad \text{eqn S2}$$

RGR is given by $d(\log(M_t))/dt$, hence eqn S1 can be differentiated by substituting for $t=t(M_{ref})$ to calculate SGR. This gives

$$SGR_i = \exp(r_i)\left(A_i - \log(M_{ref})\right) \qquad \text{eqn S3}$$

Table S1. List of the ten plant lines used in this experiment with line ID, source of the seeds and description of the phenotypes. NASC is the Arabidopsis Stock Centre (Nottingham, UK).

Line	Line ID	Source	Mechanism of knockout	Phenotype
Wildype	N1092	NASC	-	Col-0
gl1-2	CS3126	Dr. M. Ohto, UC Davis, USA	X-ray induction	No leaf trichomes on early leaves (e.g. glabrous), decreased phenolic expression
nst1-2	SALK_149993	NASC	T-DNA insertion	Knockout in *NAC Secondary Wall Thickening Promoting Factor 1*. Slight reduction in secondary wall thickening of inflorescence stems.
ppi1-2	SALK_009375	NASC	T-DNA insertion	Knockout in *Proton Pump Interactor 1*. No phenotypic differences to wildtype under a variety of biotic and abiotic stresses are reported.
myb28	BRC_H161b	Prof P. Morandini, University of Milan, Italy	T-DNA insertion	No visible phenotype reported. Reduction of short-chain aliphatic glucosinolates by ~50 % compared to wildtype. Long-chain aliphatic glucosinolates are abolished.
myb29	SM3.34316	Prof P. Morandini, University of Milan, Italy	T-DNA insertion	No visible phenotype reported. Reduction of short-chain aliphatic glucosinolates by ~50 % compared to wildtype.
myb2829	-	Prof P. Morandini, University of Milan, Italy	T-DNA insertion	Delay in seed germination and initial growth. Aliphatic glucosinolates abolished.
cyp79B2	-	Prof J. Celenza, Boston University, MA, USA	T-DNA insertion	No visible phenotype reported. Reduction of indolic glucosinolates by ~20 %.
cyp79B3	-	Prof J. Celenza, Boston University, MA, USA	T-DNA insertion	No visible phenotype reported. Reduction of indolic glucosinolates by ~40 %.
cyp79B2B3	-	Prof G. Jander, Cornell University, NY, USA	T-DNA insertion	Delay in growth. Indolic glucosinolates abolished.

Table S2. Candidate models for the asymptotic regression model of plant growth. The first line is the full model where there are interactive effects of line identity and herbivory treatment on all three model parameters (A_i, r_i and $M_{i,0}$). Stars (*) represent significant terms based on F-tests. A reduction of two AIC units is generally accepted to indicate a preferred model. Based on this criterion, the final preferred model is marked in boldface.

Asymptotic mass (A_i)	Rate constant (r_i)	Starting mass ($M_{i,0}$)	AIC
Line x Herbivory***	Line x Herbivory***	Line x Herbivory	223.91
Line x Herbivory***	Line x Herbivory**	Line + Herbivory*	216.44
Line x Herbivory***	Line x Herbivory*	Herbivory*	213.46
Line x Herbivory**	Line*** + Herbivory***	Herbivory*	204.60
Line* + Herbivory***	**Line*** + Herbivory***	**Herbivory**.	**192.25**
Line*** + Herbivory***	Line*** + Herbivory***		194.51
Line***	Line***		217.30

* P<0.05; **P<0.01; ***P<0.001

Table S3. Candidate models for the asymptotic regression model of aphid rate of reproduction. The first line is the full model where there are effects of line identity on all three model parameters (A_i, r_i and $M_{i,0}$). Stars (*) represent significant terms based on F-tests. A reduction of two AIC units is generally accepted to indicate a preferred model. Based on this criterion, the final preferred model is marked in boldface.

Asymptotic number	Rate constant	Starting number	AIC
Line***	Line***	Line	26.71
Line* **	**Line*** **		**22.12**
Line***			52.41

* P<0.05; **P<0.01; ***P<0.001

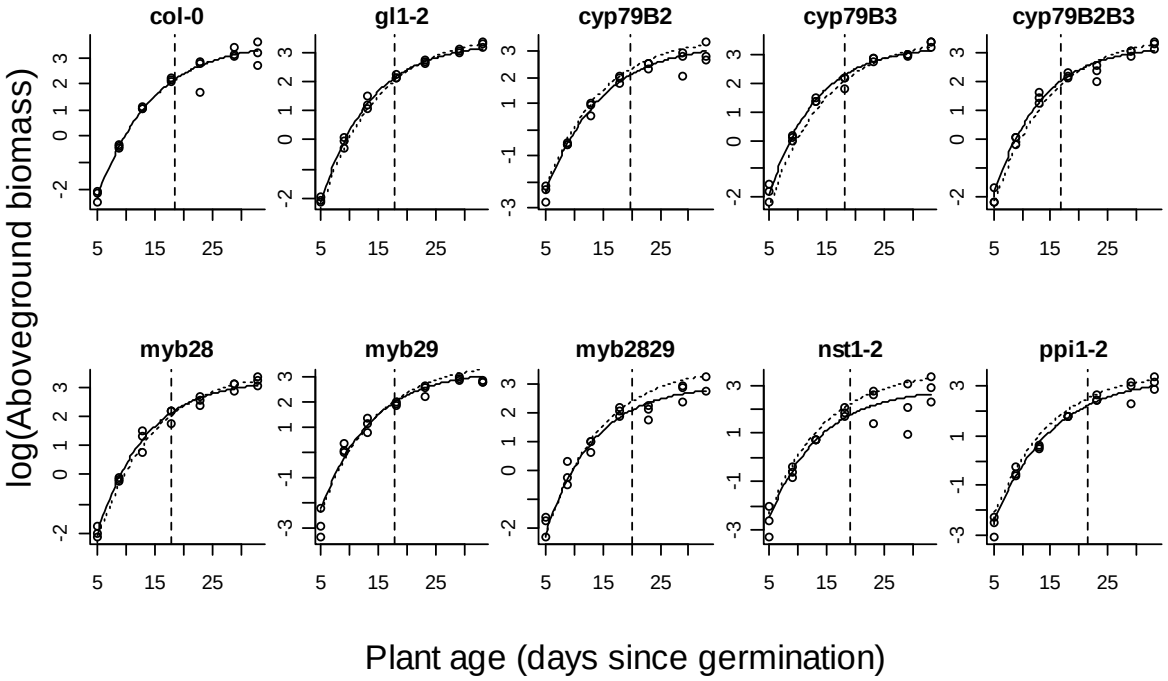


Figure S1. Data plus fitted growth curves for the ten plant lines (wildtype (col-0) plus the nine knockout mutants) in the absence of herbivory using parameters taken from the final model. On each panel the line-specific regression (solid line) is shown together with the growth curve of the wildtype (dotted line). Vertical dashed lines represent the average bolting age observed in each line.

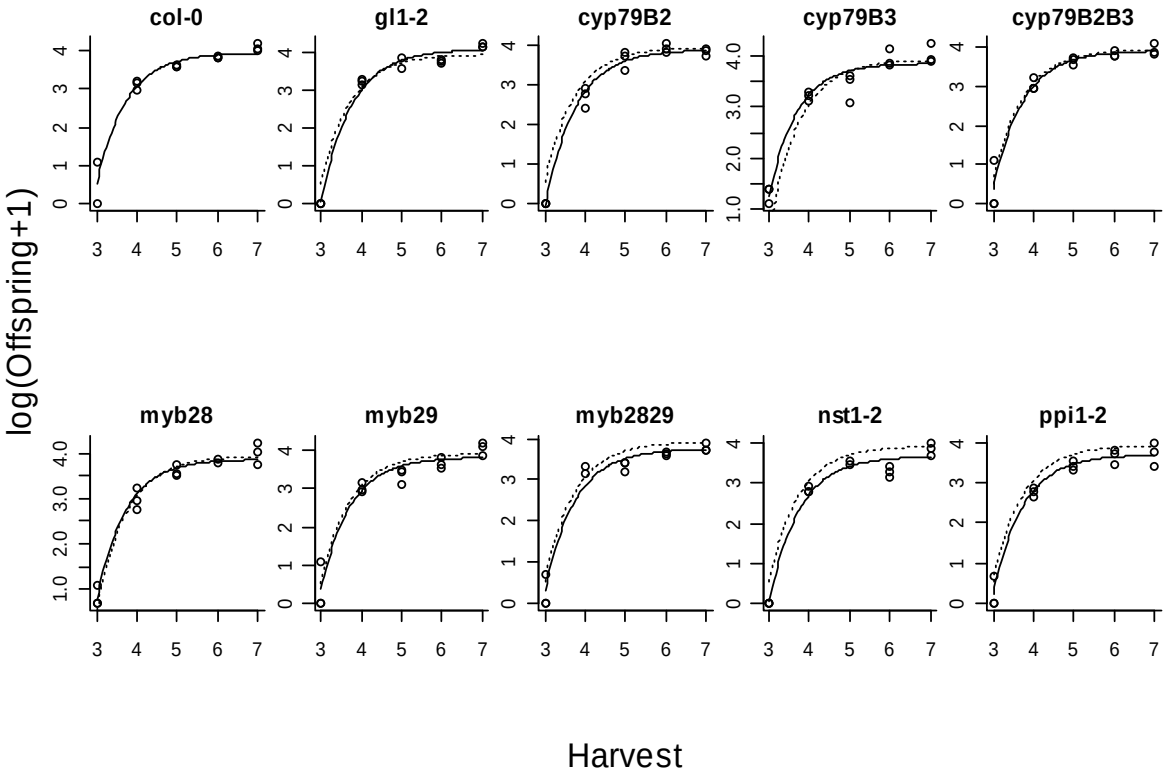


Figure S2. Data plus fitted curves for the number of F_1 aphids on the ten plant lines (wildtype (col-0) plus the nine knockout mutants) using parameters taken from the final model. On each panel the line-specific regression (solid line) is shown together with the growth curve from the wildtype (dotted line). First offspring production was noted on the third harvest date (8 days after introduction of the focal aphid).

CHAPTER 3

NON-LINEAR ROSETTE AREA MODELLING FOR A SIMPLER, MORE ROBUST QUANTIFICATION OF GENOTYPE EFFECTS

Tobias Züst, Christopher Philipson, Andy Hector, and Lindsay A. Turnbull

ABSTRACT

Premise of the study: High-throughput methods for automated phenotyping are increasingly powerful, affordable and available for academic research. Such automated methods have many advantages including the ability to measure large sample sizes with high precision while removing the need for destructive sampling. However, while such systems have the potential to uncover otherwise overlooked phenotypic patterns in plants, they also highlight the need for biologically meaningful statistical methods to analyse the very large datasets they create.

Methods: In this paper we use a recently published dataset (Arvidsson et al., 2011) generated by a camera-based plant phenotyping system to demonstrate some of the advantages of non-linear over polynomial linear analysis methods for calculating and comparing plant growth rates.

Key results: We demonstrate the following limitations of the polynomial linear model analysis: 1) the parameters have no clear biological meaning; 2) predictions are poor beyond the range of the data and 3) there is no clear underlying growth process, for example, it would be impossible to simulate new plants growing in new situations. In contrast, we show how non-linear models do not suffer from these problems and in addition allow the calculation of relative growth rates at one or more common sizes.

Conclusions: In the case examined here, comparison of relative growth rates at common sizes overturns the original conclusions of the authors and instead suggests that some of the differences revealed by the linear model analysis are unlikely to be biologically significant.

INTRODUCTION

High-throughput methods for automated phenotyping can accurately measure large numbers of plants while removing the need for destructive sampling. However, such systems highlight the need for biologically meaningful statistical methods to analyse the very large datasets they create. In this paper we use a recently published dataset of daily rosette-area measurements of *Arabidopsis thaliana* (Arvidsson et al., 2011) to demonstrate some of the advantages of non-linear regression analyses (or non-linear mixed-effects models) over linear methods for calculating and comparing growth rates.

Conventional relative growth rate (RGR) is a problematic measure of growth since RGR is usually size-dependent, giving smaller individuals higher values than large ones, even if they share an identical growth rate at a common size (Turnbull *et al.*, 2008; Rose *et al.*, 2009; Taylor *et al.*, 2010; Renton & Poorter, 2011). In the past, we, and others, have found that using non-linear growth functions and size-standardised growth rates (SGR) in place of log-linear functions and conventional RGR produces predictions that are biologically more meaningful (Rose et al., 2009; Paul-Victor et al., 2010).

Plants can only rarely sustain exponential growth (linear increases when biomass is transformed on the log-scale). As plants get larger, deviations from exponential growth can occur for a variety of reasons, including increasing allocation to structural non-photosynthetic tissue, self-shading of the leaf canopy and increasing belowground resource limitation (Evans, 1972; Ingestad & Agren, 1992; Maranon & Grubb, 1993; Enquist *et al.*, 1999). This problem is most severe if RGR is calculated from only two size measurements because deviations from a log-linear function are not detectable and hence neglected. One solution to account for non-linearity is the fitting of polynomial models in a linear model framework (Poorter, 1989). However, polynomials have the following key problems:

1) *Interpretation*: polynomial equations are purely empirical models that introduce arbitrary levels of curvature as required to fit patterns in a given dataset. It is therefore difficult to make biological interpretations of the coefficients in a polynomial model (Pinheiro & Bates, 2000; Paine *et al.*, in press).

2) *Prediction*: while polynomials, by their very nature, can potentially provide an excellent fit to the observed values, they tend to make poor predictions beyond the range of the data.

3) *Over-fitting*: polynomials do not assume any particular underlying growth process. There is therefore a risk that they will 'overfit', mistaking curvature introduced by noise for a genuine change in growth rate; in these cases the model would be expected to perform poorly when applied to new datasets. For example, analysis of a new dataset might mistakenly detect an apparently significant cubic term (a false positive). In contrast, adopting a non-linear function such as the power law constrains us to only consider patterns of growth we think are realistic based on the underlying biology.

In contrast, non-linear models assume some underlying growth process and can be seen as semi-mechanistic. This means that their coefficients are easier to interpret, that they are more likely to make better predictions and, because of the constraint imposed by the assumed underlying growth process, they are less likely to overfit and should perform better when applied to new datasets of the same type. Parameters values from different datasets can also be directly compared, e.g. if a power-law growth model (Eqn 1) is fitted to a range of different datasets, then the values of the scaling exponents can be meaningfully compared, and indeed used to test theoretical predictions (Enquist *et al.*, 1998; Enquist *et al.*, 1999; Muller-Landau *et al.*, 2006; Coomes & Allen, 2009).

Here, we use the raw data provided by Arvidsson et al. (2011) in their online supplementary material to demonstrate how non-linear modelling is a simple alternative to the

linear analysis proposed by the authors. R code is provided as online supplementary material (Appendix 1).

MATERIALS AND METHODS

The first data set in Arvidsson et al. (2011) was collected on two genotypes: the wildtype *col-0* and the starch degradation mutant *sex4-3*, in order to compare their growth rates. Arvidsson et al. 2011 developed a camera-based plant phenotyping system which takes daily photographs of plants and automatically measures the area and shape parameters of rosette leaves. They also developed an analysis pipeline, which automatically plots graphs and fits linear mixed-effects models to log-transformed rosette areas and calculates relative growth rates (RGR).

Arvidsson et al. (2011) modelled log-transformed rosette area as a function of age using a linear mixed-effects model, with fixed effects of genotype and light intensity measured at the individual plant level. To deal with the deviation from a log-linear growth function, they included a quadratic age term. Rosette growth in *Arabidopsis* might be expected to deviate from log-linear for two different biological reasons. First, during the vegetative phase of plant growth, exponential growth may be impossible to sustain for the reasons outlined above. Second, as the maximum rosette size is approached, the plant switches to reproduction and future growth is invested directly in the inflorescence. Ecological theory suggests that the switch between vegetative and reproductive growth should be rather abrupt and absolute (Cohen, 1976), but little experimental data is available on the efficiency of this switch and some slowing of growth is likely to occur before bolting is observed.

The full data set includes rosette areas up to 39 days after germination (hereafter age) but contains no information on bolting day. We fitted a range of non-linear models, including a monomolecular and a 4-parameter logistic growth model (Paine et al., in press), but it was clear that there were insufficient measurements at the later growth stages for good estimates of the asymptotic size. Instead we present a power-law model fit to the cropped data to demonstrate our general approach. Close examination of the log-linear model with a quadratic term advocated by Arvidson et al. (2011) reveals that for approximately the last eight days, the model predicts larger sizes than were actually observed. The most likely explanation for this overprediction is the switch from vegetative to reproductive growth towards the end of the experiment. A similar overprediction occurred with the power-law as would be expected given that it only applies to the vegetative growth phase; hence we cropped the dataset to include all data measured up to day 26 for *col-0* and up to and including day 30 for *sex4-3*. The mutant *sex4-3* clearly grows more slowly than *col-0*, but on day 30 reaches a size comparable to the size of *col-0* on day 26. In such a case of clearly different lengths of the vegetative growth phase, genotype-specific cropping of the data is most appropriate for direct comparison of genotypes, unless more data was available that would allow inclusion of an asymptotic term.

The power-law function (Enquist et al., 1998; Enquist et al., 1999), assumes that absolute growth rate is proportional to current mass raised to some power:

$$\frac{dM}{dt} = \alpha M^{\beta} \qquad \text{(Eq. 1)}$$

where α is the growth coefficient and β is the scaling exponent. The size-standardised growth rate or SGR is given by:

$$SGR = \frac{1}{M}\frac{dM}{dt} = \alpha M_c^{\beta-1} \qquad \text{(Eq. 2)}$$

where M_c is a common reference size (Paine et al., in press). On the log-scale, this gives:

$$\log(SGR) = \log(\alpha) + (\beta - 1)\log(M_c) \quad \text{(Eq. 3)}$$

The two parameters of the power-law function thus have clear biological interpretations: α, the growth coefficient determines the maximum growth rate in the absence of size-dependent growth limitation and β, the scaling exponent, describes how rapidly growth rate declines with increasing size, i.e. it is a measure of size-related growth inefficiency. We might therefore expect these parameters to be influenced by different genetic and environmental factors. For example, the scaling exponent β might be affected by leaf shape, because different leaf shapes might lead to different degrees of overlap and hence self-shading within the rosette. In contrast, the growth coefficient α is likely to increase with increasing nutrient availability, but in addition will be affected by intrinsic genetic factors such as allocation to costly defensive compounds (Paul-Victor et al., 2010; Züst et al., 2011). In contrast, the parameters of the quadratic equation cannot be so easily decomposed. Because we expect some systematic deviations from the power-law as the plant approaches the switch to the reproductive growth phase, a power-law should only be fitted to the vegetative growth phase, while data for the few days prior to bolting should probably be excluded.

Using the data from the first experiment in Arvidsson et al. (2011), we modelled untransformed rosette area as a function of plant age (days after germination) by fitting the closed-form solution of equation 1 when $\beta < 1$ in the nlme package for the statistical software R (R Development Core Team, 2010; Philipson et al., in press):

$$Area_t = \left(Area_0^{1-\beta} + \alpha(1-\beta)t\right)^{1/1-\beta} \quad \text{(Eq. 4)}$$

The rosette area of a plant at time t is thus a function of $Area_0$ (estimated rosette area at time = 0), the growth coefficient α, and the scaling exponent β. We treated plant identity as a random effect. Starting from a model with no fixed effects, we included effects of genotype and light

intensity on all three model parameters and compared model fits based on Akaike's Information Criterion (AIC) where a difference of more than 2 units was taken as indicating a better model.

RESULTS AND DISCUSSION

The most parsimonious model according to the AIC supported a genotype effect on the initial leaf area $Area_0$, the growth coefficient α, and the scaling exponent β, and a positive effect of light intensity on the growth coefficient α. No simplification of the random effects was possible; i.e. the final model included an effect of plant identity on all three parameters. As we modelled the untransformed data, the residuals were heteroscedastic. However, this could be accounted for by including a *varExp* variance structure, which multiplies the variance σ^2 with an exponential function of age and an additional parameter δ (Zuur et al., 2009). Compared to the wildtype, the mutant had lower values of the initial leaf area, $Area_0$ and of the growth coefficient, α, but a higher value of the scaling exponent, β (Table 1).

The time-based RGR comparison carried out by Arvidsson et al. (2011) shows that the mutant *sex4-3* grows more slowly initially, but that later in the experiment it overtakes and grows faster than the wild type. This is most likely an artefact of the different range in the vegetative growth phase of the two genotypes, and a size-based comparison allows us to see this clearly. Using estimated model parameters and equation 3, we calculated SGR for the two *Arabidopsis* genotypes *col-0* and *sex4-3* over the whole range of observed sizes (Fig 2). For the power-law fit, SGR does not depend on the initial size of a plant, but instead compares growth rates at a given size; hence SGR is not confounded with germination day or seed weight. A higher value of the scaling exponent β results in a slower decline of SGR with size (eqn 3), hence the difference in SGR between the mutant and the wildtype decreases with

increasing leaf area – appearing to confirm that the mutant will indeed overtake the wildtype at some size (Fig 2). However, the lower value of the growth coefficient α means that the mutant grows more slowly than the wildtype at all observable sizes. In fact, the SGR function of *sex4-3* is predicted to cross with *col-0* at the implausible reference size of 269 cm^2 while even in the full dataset the maximum size achieved by a plant was 23 cm^2! Thus while the difference in scaling exponents of the two genotypes is statistically significant and improves the overall model fit, it has only low biological relevance, an important distinction.

The key advantage of non-linear modelling in this case is the improved biological interpretation. This results from assuming a likely growth function, instead of just empirically fitting an arbitrary polynomial function to a set of data points. This means that non-linear models are likely to extrapolate in a more appropriate way when making predictions and transfer more reliably to new datasets of a similar type. In general, an important potential benefit of non-linear modelling is the use of fewer parameters. While that is not the case in this analysis (the power law requires as many parameters as the quadratic model) it would be an advantage compared to higher-order polynomials or other complex models. Non-linear regression may also provide a better fit to the data than linear models with polynomials, although once again that potential advantage is not important here as both models explain a very large amount of the variability in the data.

In conclusion, we recommend that non-linear modelling is a more biologically informative way of estimating aspects of the vegetative growth phase of *Arabidopsis*. In particular the power-law model provides an excellent fit to the early vegetative growth phase and can be implemented using non-linear regression or non-linear mixed-effects models (as appropriate). Fitting a power-law function to the vegetative growth phase of *Arabidopsis* yields estimates of three parameters: the initial size on day 0 (which is likely to be closely related to seed size), the growth coefficient (determining the maximum growth rate in the

absence of size-dependent growth limitations), and the scaling exponent (determining the diminishing returns in the size-inefficiency of growth). In this application, the starch-degradation mutant had significantly different values for three of the measured parameters when compared to the wildtype, but the biologically relevant difference in growth rate was due only to the decreased growth coefficient α in the mutant. Over the whole range of biologically reasonable sizes, this decreased growth coefficient led to a reduction in SGR, hence the starch-degradation mutant can never catch up or overtake the wildtype, as suggested by the original linear-model analysis.

LITERATURE CITED

Arvidsson S, Pérez-Rodríguez P, Mueller-Roeber B. 2011. A Growth Phenotyping Pipeline for *Arabidopsis Thaliana* Integrating Image Analysis and Rosette Area Modeling for Robust Quantification of Genotype Effects. *New Phytologist* **191**(3): 895-907.

Cohen D. 1976. Optimal Timing of Reproduction. *American Naturalist* **110**(975): 801-807.

Coomes DA, Allen RB. 2009. Testing the Metabolic Scaling Theory of Tree Growth. *Journal of Ecology* **97**(6): 1369-1373.

Enquist BJ, Brown JH, West GB. 1998. Allometric Scaling of Plant Energetics and Population Density. *Nature* **395**(6698): 163-165.

Enquist BJ, West GB, Charnov EL, Brown JH. 1999. Allometric Scaling of Production and Life-History Variation in Vascular Plants. *Nature* **401**(6756): 907-911.

Evans GC. 1972. *The Quantitative Analysis of Plant Growth*. Oxford: Blackwell Scientific Publications.

Ingestad T, Agren GI. 1992. Theories and Methods on Plant Nutrition and Growth. *Physiologia Plantarum* **84**(1): 177-184.

Maranon T, Grubb PJ. 1993. Physiological-Basis and Ecological Significance of the Seed Size and Relative Growth-Rate Relationship in Mediterranean Annuals. *Functional Ecology* **7**(5): 591-599.

Muller-Landau HC, Condit RS, Chave J, Thomas SC, Bohlman SA, Bunyavejchewin S, Davies S, Foster R, Gunatilleke S, Gunatilleke N, Harms KE, Hart T, Hubbell SP, Itoh A, Kassim AR, LaFrankie JV, Lee HS, Losos E, Makana JR, Ohkubo T, Sukumar R, Sun IF, Supardi NMN, Tan S, Thompson J, Valencia R, Munoz GV, Wills C, Yamakura T, Chuyong G, Dattaraja HS, Esufali S, Hall P, Hernandez C, Kenfack D, Kiratiprayoon S, Suresh HS, Thomas D, Vallejo MI, Ashton P. 2006. Testing Metabolic Ecology Theory for Allometric Scaling of Tree Size, Growth and Mortality in Tropical Forests. *Ecology Letters* **9**(5): 575-588.

Paine CET, Marthews TR, Vogt DR, Purves D, Rees M, Hector A, Turnbull LA. in press. How to Fit Nonlinear Plant Growth Models and Calculate Growth Rates: An Update for Ecologists. *Methods in Ecology and Evolution*.

Paul-Victor C, Züst T, Rees M, Kliebenstein DJ, Turnbull LA. 2010. A New Method for Measuring Relative Growth Rate Can Uncover the Costs of Defensive Compounds in *Arabidopsis Thaliana*. *New Phytologist* **187**(4): 1102-1111.

Philipson CD, Saner P, Marthews TR, Nilus R, Reynolds G, Turnbull LA, Hector A. in press. Light-Based Regeneration Niches: Evidence from 21 Dipterocarp Species Using Size-Specific Rgrs. *Biotropica*.

Pinheiro JC, Bates DM. 2000. *Mixed-Effects Models in S and S-Plus*. London: Springer.

Poorter H. 1989. Plant Growth Analysis: Towards a Synthesis of the Classical and the Functional Approach. *Physiologia Plantarum* **75**(2): 237-244.

R.Development.Core.Team 2010. R: A Language and Environment for Statistical Computing.In. Vienna, Austria: R Foundation for Statistical Computing. R Development Core Team.

Renton M, Poorter H. 2011. Using Log-Log Scaling Slope Analysis for Determining the Contributions to Variability in Biological Variables Such as Leaf Mass Per Area: Why It Works, When It Works and How It Can Be Extended. *New Phytologist* **190**(1): 5-8.

Rose KE, Atkinson RL, Turnbull LA, Rees M. 2009. The Costs and Benefits of Fast Living. *Ecology Letters* **12**: 1379-1384.

Taylor SH, Hulme SP, Rees M, Ripley BS, Woodward FI, Osborne CP. 2010. Ecophysiological Traits in C_3 and C_4 Grasses: A Phylogenetically Controlled Screening Experiment. *New Phytologist* **185**(3): 780-791.

Turnbull LA, Paul-Victor C, Schmid B, Purves DW. 2008. Growth Rates, Seed Size, and Physiology: Do Small-Seeded Species Really Grow Faster? *Ecology* **89**(5): 1352-1363.

Züst T, Joseph B, Shimizu KK, Kliebenstein DJ, Turnbull LA. 2011. Using Knockout Mutants to Reveal the Costs of Defensive Traits. *Proceedings of the Royal Society B: Biological Sciences* **278**: 2598-2603.

Zuur AF, Ieno EN, Walker NJ, Saveliev AA, Smith GM 2009. Mixed Effects Models and Extensions in Ecology with R.In. New York, USA: Springer Science+Media.

TABLES

TABLE 1. Parameter estimates and 95 % confidence intervals from the most parsimonious model based on AIC. This model included an additive linear effect of light intensity on the growth coefficient α (see Figure 1).

Parameter	Genotype	Lower 95% CI	Estimate	Upper 95% CI
$Area_0$	*col-0*	0.00510	0.00547	0.00587
$Area_0$	*sex4-3*	0.00433	0.00472	0.00513
A	*col-0*	0.190	0.199	0.208
A	*sex4-3*	0.153	0.163	0.173
B	*col-0*	0.892	0.896	0.900
B	*sex4-3*	0.913	0.918	0.923

FIGURES

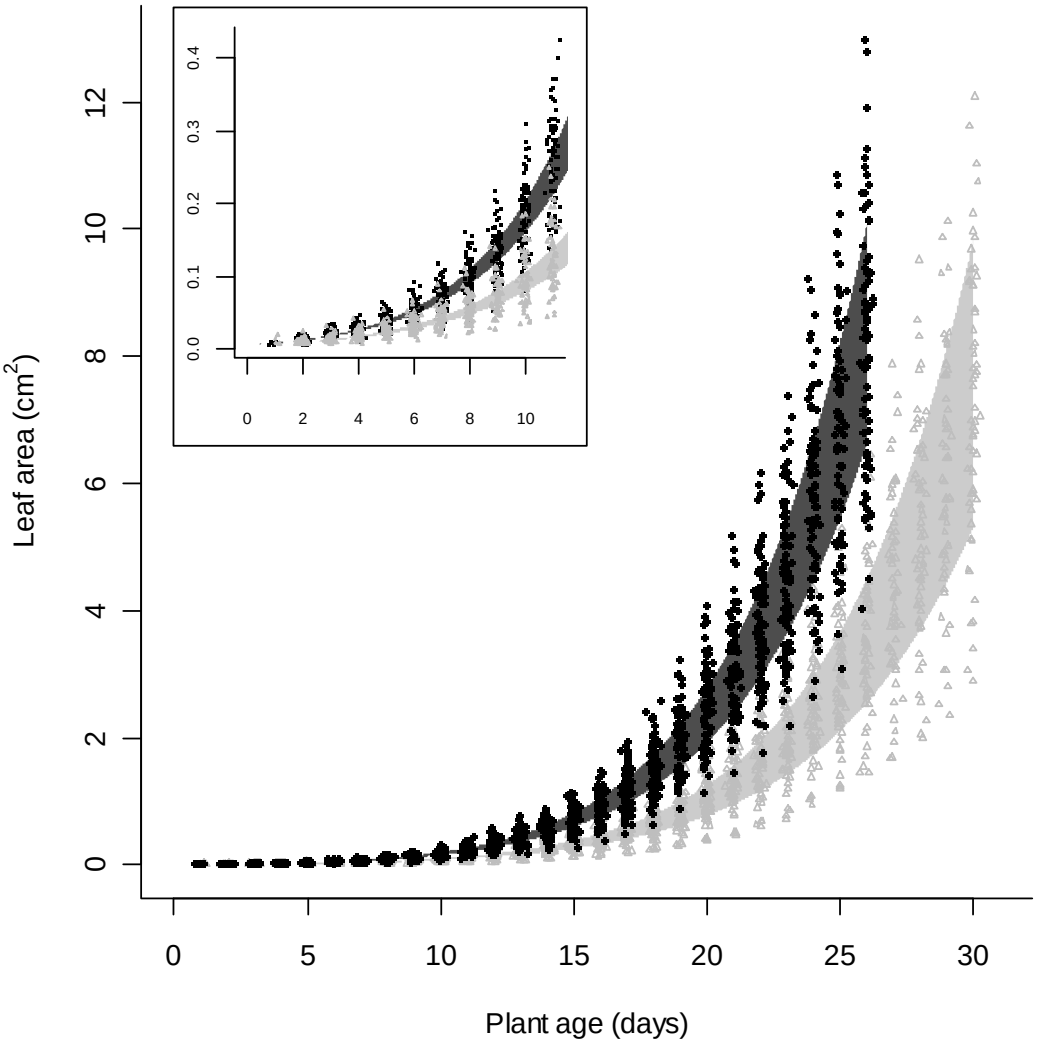


FIGURE 1. Power-law fit to rosette leaf area as a function of plant age (days after germination) for the wildtype *col-0* (black circles) up to day 26 and the mutant *sex4-3* (grey triangles) up to day 30. The filled areas represent the mean model fit for the full range of light intensities measured at the plant level for *col-0* (black) and *sex4-3* (grey). Datapoints are jittered slightly for clarity. The inset is a magnification of the curve for the first 11 days of growth.

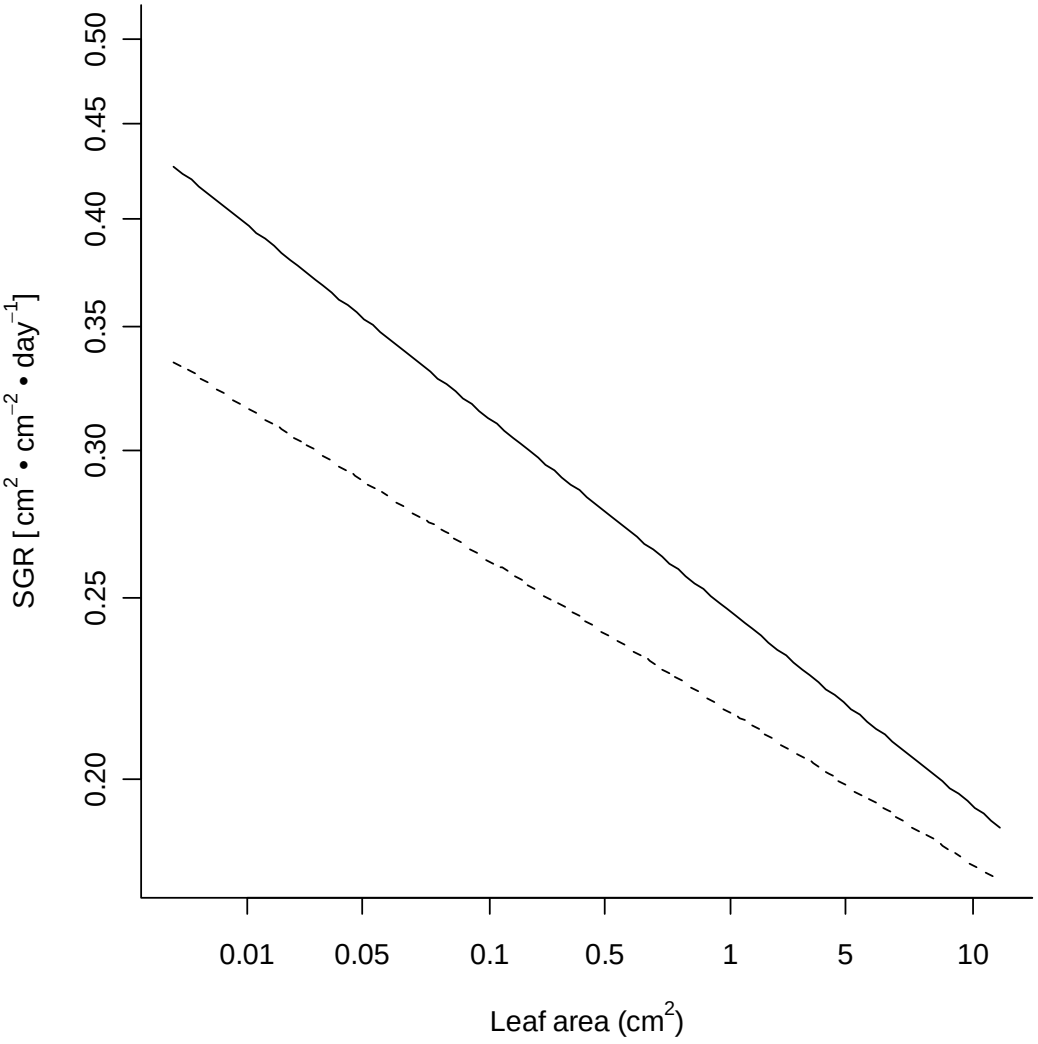


FIGURE 2. Predicted size-standardised growth rate (SGR) as a function of leaf area for both the wildtype *col-0* (solid line) and the mutant *sex4-3* (dashed line) at mean light intensity. The range of the x-axis spans the range observed in plants of *col-0* in experiment 1 up to and including day 26.

CHAPTER 4

QUANTIFICATION OF GROWTH RATES AND HERBIVORY ACROSS RIL POPULATION OF *ARABIDOPSIS THALIANA* REVEALS THE COSTS OF AN EFFECTIVE PATHWAY OF PLANT DEFENCE

Tobias Züst, Majid Iravani, Gabriela Schaepman-Strub, Jason A. Corwin, Baohua Li, Bindu Joseph, Daniel J. Kliebenstein and Lindsay A. Turnbull

SUMMARY

Understanding the trade-offs between growth rates, chemical defence and herbivore damage is of fundamental and applied importance. However, carrying out such studies on large populations requires efficient quantification of whole plant growth rates using non-destructive techniques. We present a high-accuracy, low-cost method for measuring rosette growth rates of *Arabidopsis thaliana* and apply this method to the complete Kas x Tsu population of recombinant-inbred lines (RILs). We also measured the concentration of fourteen glucosinolate compounds in leaves of all lines and conducted herbivore trials with two caterpillar species on a subset of these lines. In contrast to work on other genotypes of *A. thaliana*, we found only weak correlations between growth rate, glucosinolate content, and herbivore damage. We hypothesised that this difference is caused by the complete lack of aliphatic glucosinolates with a 4-carbon side-chain (4C) within the Kas x Tsu population. To test this, we directly compared the growth rates of selected lines from the Kas x Tsu population with those from the L*er* x Cvi population. As predicted, the L*er* x Cvi population grew more slowly than the Kas x Tsu population, and a significant fraction of this difference could be attributed to the presence of the *4C*-allele at the *GS-ELONG* locus in the L*er* x Cvi population. Our work provides strong evidence for the high cost of 4C aliphatic glucosinolates which appear to be the most effective defensive compounds against generalist, leaf-chewing herbivores.

INTRODUCTION

Plants employ a wide range of defensive traits to deter and avoid damage by herbivores. Defence traits vary in their effectiveness against different herbivore species, and these trait-for-herbivore interactions create the basis for community-wide co-evolutionary arms races

which are a major source of biodiversity (Thompson, 1999). A major assumption of plant defence theory is that defensive traits have some cost to the plant, i.e., that they trade off with fitness (Coley *et al.*, 1985; Bazzaz *et al.*, 1987; Herms & Mattson, 1992). While the seed set of a plant is often considered to be the best estimate of lifetime fitness, plant growth rate can also be a good predictor of the outcome of competition and is thus a useful surrogate for fitness in competitive environments (Fakheran et al., 2010; Züst et al., 2011).

Traditionally, relative growth rate or RGR is considered to be the best measure of plant growth rate. RGR is typically based on only two biomass measurements and is calculated using:

$$RGR = \frac{\log(M_2 / M_1)}{t_2 - t_1} \qquad \text{Eqn 1}$$

However, these traditional methods for calculating RGR implicitly assume that plants grow exponentially (linear on the log-scale), in which case RGR would be constant and unbiased by size. In fact, as plants get larger, deviations from exponential growth occur for a variety of reasons, including increasing allocation to structural non-photosynthetic tissue, self-shading of the leaf canopy and increasing below-ground resource limitation (Evans, 1972; Ingestad & Agren, 1992; Maranon & Grubb, 1993; Enquist *et al.*, 1999). As a solution, we and other authors have repeatedly advocated the use of non-linear growth functions that allow the calculation of size-standardized RGR (called SGR) and facilitate comparison among species or genotypes (Rose et al., 2009; Paul-Victor et al., 2010; Paine et al., in press).

While non-linear functions can capture growth processes more accurately, they require more data collection. As a consequence, practical methods for non-destructive, high-resolution plant size measurements are in demand, especially in the light of large-scale attempts to map growth-related traits to the genome. Recently, high-throughput methods for automated growth phenotyping have become available for academic research (e.g., Durham

Brooks et al., 2010; Arvidsson et al., 2011). However, these methods employ custom-built robotic systems that are only affordable to a small proportion of the scientific community. Here we present a new, high-accuracy method for measuring rosette leaf area that uses conventional, man-made photographs and a software package that is commonly used for remote sensing applications. This method is applicable to all two-dimensional growth processes, and it is well-suited to describe the rosette growth phase of the model plant *Arabidopsis thaliana* L. (Brassicaceae).

A. thaliana employs a set of defence-related traits: most importantly a range of secondary metabolites belonging to the glucosinolates with known defensive properties against herbivores (Bones & Rossiter, 1996). Glucosinolates consist of a sulphur-linked glycone moiety, a nitrogen-linked sulphate, and a variable side chain (Mithen et al., 1995). This side-chain is the biologically active part and may contain aliphatic, indolyl, or aromatic groups. At least 43 different glucosinolate compounds are present in *A. thaliana* (Reichelt et al., 2002; Kliebenstein et al., 2007), and the majority of this variation has been explained by the combination of functional and non-functional (null) alleles at four genetic loci (Mithen et al., 1995; Kliebenstein et al., 2001b). The combinations of alleles at these loci are responsible for the formation of distinct glucosinolate profiles, or 'chemotypes'. There are six major chemotypes, resulting from the combination of *GS-ELONG*, which regulates the carbon side-chain elongation of aliphatic glucosinolates (either 3-carbon (*3C*) or 4-carbon (*4C*)), and *GS-AOP*, which controls the conversion of methylsulfinylalkyl (*NULL*) to either alkenyl (*ALK*) or hydroxypropyl (*OHP*) glucosinolates. Some of these chemotypes can be further modified by *GS-OH*, responsible for the conversion of 3-butenyl to 2-hydroxy-3-butenyl glucosinolate; or *GS-OX*, which regulates the conversion of methylthioalkyl to methylsulfinylalkyl glucosinolates. The chemotype of *A. thaliana* also plays an important role in determining the respective herbivore community (Bidart-Bouzat & Kliebenstein, 2008), and some glucosinolate compounds have clear deterrent properties against particular herbivore species

(Lambrix et al., 2001). Glucosinolates are generally considered to incur costs to the plant, and glucosinolate concentrations tend to be negatively correlated with seed set (Mauricio, 1998) and growth rate (Paul-Victor et al., 2010). However, different glucosinolate compounds within a chemotype are usually strongly intercorrelated, making it difficult to quantify the relative importance of individual compounds.

We measured the glucosinolate content and rosette growth rates of a complete RIL population consisting of 341 inbred lines (McKay et al., 2008), derived from a cross between the two accessions Kashmir (Kas-1, N903) and Tsushima (Tsu-1, N1640). Kas-1 has a *3C-ALK* chemotype, while Tsu-1 has a *3C-OHP* chemotype. Thus, both parents carry the *3C-* allele at the *GS-ELONG* locus and the resulting RIL population consists of a roughly equal mixture of only two chemotypes (*3C-ALK* and *3C-OHP*). We measured damage inflicted by a generalist and a specialist herbivore on a subset of all lines, and correlated growth rates, glucosinolate content and herbivore damage in a search for growth/defence trade-offs. We then compared these correlations with a recent study (Paul-Victor et al., 2010), where we had used a RIL population derived from a cross between the *A. thaliana* accessions Landsberg *erecta* (L*er*, N8581) and Cape Verdi Islands (Cvi-1, N8580, Alonso-Blanco et al., 1998). L*er* has a *3C-OHP* chemotype whereas Cvi-1 has a *4C-ALK* chemotype. The presence of the *4C-* allele at the *GS-ELONG* locus in one of the parent lines means that the L*er* x Cvi population consists of four chemotypes: *3C-ALK*, *3C-OHP*, *4C-ALK* and *4C-OHP*. The comparison of these two RIL populations thus provides a direct test of the effect of the *GS-ELONG* locus and the production of *4C*-glucosinolate compounds on growth rate and herbivore resistance. Such clean tests of specific loci and compounds are otherwise prevented by intercorrelations among glucosinolate compounds.

MATERIALS AND METHODS

Growth of the Kas x Tsu RIL Population. Seeds of the 341 lines of the Kas x Tsu recombinant inbred population were obtained from the Arabidopsis Biological Resource Center (ABRC, Columbus OH, USA). We grew a total of four to five plants per line, split into two blocks separated by approximately three months. We grew plants in large planting trays with 104 individual wells (30 x 25 x 100 mm), filled with standard potting soil (Sunshine Mix #1, Sun Gro Horticulture, Bellevue WA, USA). Prior to sowing, we imbibed seeds in distilled water and cold stratified them at 4° C for four days. We placed approximately 3-5 seeds of a single genotype in the centre of a well and covered the trays with a transparent plastic hood to retain humidity during germination. We randomized genotype placement within the trays to create a randomized complete block design split between two blocks. We recorded the germination day for all plants. After one week, we removed the transparent hoods and surplus plants to leave one seedling per well. We watered plants twice a week with nutrient-enriched water, using 0.5 % N:P:K fertilizer in a 2:1:2 ratio (Grow More 4-18-38, Grow More Inc., Gardena CA, USA) and kept them in a climate-controlled chamber at 22° C and a day/night cycle of 10h/14h.

To record plant growth, we photographed each individual planting tray every second day, starting from day 7 after sowing. We photographed the first block of the experiment in TIFF format using a Nikon Coolpix E995 digital camera, and the second block in Jpeg format using a Fujifilm FinePix S1500 digital camera. All pictures were taken by the same technician, holding the camera at breast height. Flash was disabled and we took the pictures under ambient (artificial) light with automatic camera settings. Thirty-one days after sowing, we harvested plants for glucosinolate analysis. We removed two leaves from the first fully mature leaf pair from each plant and stored leaves in 90% methanol to inhibit enzymatic breakdown of chemical compounds. We extracted glucosinolates and analysed them by HPLC

according to previously described methods (Kliebenstein et al., 2001a; Kliebenstein et al., 2001b).

Image Processing and Extraction of Individual Leaf Areas. We extracted the rosette area of plants from photographs using the image analysis software ENVI (Version 4.6, ITT Vis, Boulder, USA) and supervised maximum likelihood classification, which uses digital number values of the images to separate plant area from the soil background (Fig. 1a). Because the distance of the camera to the trays differed slightly between photographs, the pixel/mm relation differed among images, making direct comparisons of leaf area impossible. In a first step we thus registered all images to the scale of one reference image, using the image-to-image registration method as implemented in ENVI. This method requires the user to mark a minimum of four points or 'nodes' that are present in both the reference and the sample image. The sample image is then transformed in a way that all nodes of the sample image match the reference nodes. We registered all images to one reference image, placing up to eight nodes at the corners of corresponding wells within the planting trays. In the next step, we classified the images into three categories, namely *leaf*, *soil* and *planting tray* using a maximum likelihood classifier in ENVI. We therefore selected training areas for the three classes in each image to calculate training class statistics based on the digital number values of the red, green, and blue layer of the image. (Fig. 1b). Based on this training sample, every pixel within an image was assigned to one of the three classes by the maximum likelihood algorithm (Fig. 1c). As the final step, we marked the area of each well of a planting tray, assigning leaf area within a well a unique plant or rosette identity. The individual plant areas were then defined as the number of *leaf* pixels within each of the 104 masked wells. A small ruler had been placed onto each planting tray that was used to transform pixel numbers into rosette areas (mm^2).

Area-based growth measures are problematic for later stages of plant growth, as leaf overlap with neighbouring plants increases. In addition, such methods can only cover the vegetative phase of growth. We were able to extract rosette areas up to 23 days after sowing by retracing leaf shapes of overlapping leaves by hand. After day 23, overlap became too large and made this method unfeasible.

Size-Standardised Growth Rates. We found that a power-law function (Enquist et al., 1998; Enquist et al., 1999) provided a good fit to the vegetative growth phase. Power-law growth assumes that the absolute growth rate is given by:

$$\frac{dM}{dt} = \alpha M^{\beta} \qquad \text{Eqn 2}$$

where M is a measure of size, α is the growth coefficient and β is the scaling exponent. We modelled untransformed rosette area as a function of days after germination, using the closed form solution of equation 2:

$$Area_t = \begin{cases} (Area_0^{1-\beta} + \alpha(1-\beta)t)^{1/1-\beta} & if \ \ \beta \neq 1 \\ Area_0 \exp^{\alpha t} & if \ \ \beta = 1 \end{cases} \qquad \text{Eqn 3}$$

The rosette area of a plant at time t is thus a function of the three parameters $Area_0$ (estimated rosette area at time = 0), the growth coefficient α, and the scaling exponent β. Equation 3 can be fitted to data in the nlme package for the statistical software R (R Development Core Team, 2010) and all three parameters can be estimated (Philipson et al., in press). We started with a model that only contained random effects of plant identity on all three parameters, essentially fitting unique growth curves to each individual plant. We then modified this model by adding fixed effects of RIL and experimental block to the model parameters and compared alternative models with different fixed effects using likelihood ratio

tests. We extracted RIL-specific parameters from the best model and calculated size-standardised growth rates (SGR) for each RIL. SGR is given by:

$$SGR = \alpha Area_{ref}^{\beta-1} \quad \text{Eqn 4}$$

where $Area_{ref}$ is a common size at which growth rate is calculated (Paine et al., in press).

Herbivory Assay. In order to relate growth rate and chemical defences to herbivore resistance, we measured herbivore damage caused by larvae of two lepidopteran species on a subset of the Kas x Tsu RIL population. We selected a subset of the population by ranking lines within each chemotype (*3C-OHP* and *3C-ALK*) according to their growth rate and assigning them to 35 quantiles. We used symmetrically decreasing quantile sizes towards the tails of the growth rate distribution to increase the number of extreme growth phenotypes. From each quantile we randomly selected one line for the herbivory trials, resulting in a total of 70 lines. We grew five plants per line in 96-well planting trays (well size 35 x 35 x 80 mm) filled with germination soil (GO M1, Tref Group, The Netherlands). Five seeds were planted in each well and whole trays were cold stratified at 4°C for two days. Plants were grown in a controlled climate chamber at 20° C and a day/night cycle of 16h/8h and six days after sowing surplus plants were removed.

Twenty-two days after sowing, we removed the 3rd leaf pair of each plant and placed each leaf individually into a small petri dish (Ø 35mm), marked with a unique number. Petri dishes contained a solid layer of 1.5 ml Agar (1.3%) to keep leaves fresh during the trial. We added either a single 2nd instar larvae of the generalist caterpillar *Spodoptera littoralis* or the specialist *Plutella xylostella* to each petri dish and placed them back into the same climate chamber for 24 hours. We photographed leaves before and after herbivore addition in order to calculate the leaf area removed. Photographs were taken with a Canon EOS 350D digital

camera fitted with a 60 mm macro lens. We mounted the camera on a stand with an integrated diffuse light source in the base. Petri dishes were placed on top of the light source and photographed from above, thus maximizing contrast between leaf and transparent agar. Leaf area before and after herbivory was measured using the open-source image processing software ImageJ (Rasband, 1997-2009). Using batch mode, all images were transformed into black-and-white colours and the number of black pixels was recorded in each picture, representing leaf area. In a subset of five images, we measured pixel size of the petri dish and used the mean pixel number to calculate leaf areas in mm^2. For each leaf, we calculated the percentage area removed by the herbivore. These herbivory scores were correlated with glucosinolate concentrations measured on the same lines.

Growth Rate Comparison of Two RIL Populations. To compare results from the Kas x Tsu RIL population with our previous work on the L*er* x Cvi RIL population (Kliebenstein et al., 2002; Paul-Victor et al., 2010), we carried out a small growth experiment to directly compare growth rates of the two RIL populations when grown in the same environment. We randomly selected 30 lines from each RIL population, selecting 15 lines each from the *ALK* and the *OHP* chemotypes. Seeds of the L*er* x Cvi RIL population had originally been obtained from the European Arabidopsis Stock Center (Nottingham, UK) and had since been propagated for several generations in the lab of Prof. Ueli Grossniklaus, University of Zürich.

We grew five plants per line in 77-well planting trays (well size: 35 x 35 x 55 mm) filled with germination soil (GO M1, Tref Group, The Netherlands). Five seeds were planted in each well and whole trays were cold stratified at 4°C for two days. Plants were grown in a controlled climate chamber at 20° C and a day/night cycle of 16h/8h and six days after sowing surplus plants were removed. To record plant growth, we took a photograph of each planting tray every second day, starting from day 7 after sowing. Photographs were taken with a Canon

EOS 350D digital camera fitted with an 18 mm wide-angle lens. The camera was mounted on a stand with two integrated light sources on each side of the camera. On the base of the stand we fitted a guide rail that fitted the size of the planting trays, ensuring identical positioning of the trays on each photograph. Compared to the growth experiment of the full Kas x Tsu RIL population, plants grew much faster under the long-day conditions employed here. We thus had to stop the experiment when plants started to bolt around day 19 after germination. Leaf areas of individual plants were extracted using ENVI according to the same protocol as described above. Image-to-image registration was not necessary in this experiment as the standardized camera setup ensured identical pixel/mm relations in all images. We could only extract leaf areas up to day 17 after germination, as leaf overlap became too large for accurate area determination after this day.

We fitted a similar three-parameter power-law model to this data as for the full Kas x Tsu RIL population (eqn 2), starting with a model only containing random effects of plant identity. We fitted RIL population, *GS-AOP* allele, or *GS-ELONG* allele as fixed effects to all three parameters of the power-law model. In addition, we also fitted models with a combination of RIL population and one of the two glucosinolate loci. We judged inclusion of fixed effects based on F-tests and selected the best model based on AIC values (Table 1). Using the parameters from this model, we then calculated average population SGR values. As reference sizes we used a range of values starting from the smallest measured rosette area up to the mean rosette area of the population reaching smaller sizes on day 17. In addition, we estimated the confidence intervals around these mean values by generating population prediction intervals (Bolker, 2008; Züst et al., 2011; Paine et al., in press). The method assumes that the distribution of the parameters is multivariate normal with a variance-covariance matrix given by the inverse of the information matrix. We used the function *mvrnom*, which selects multivariate normal random deviates, and the variance-covariance matrix given by the function *vcov*. At each value of the reference size, we generated 1000 sets

of parameters to calculate a distribution of SGR values for each population. The lower and upper 95% quantile of these distributions are the boundaries of the prediction intervals.

RESULTS

Growth Rate and Secondary Metabolites of the Kas x Tsu RIL Population. Despite no obvious differences in environmental conditions, plant growth rates differed significantly between the two experimental blocks (Fig. 2a). The best power-law model included a fixed effect of block on all three parameters: the estimated rosette area at time = 0, $Area_0$; the growth coefficient, α; and the scaling exponent, β. There was a significant RIL effect on $Area_0$ and α only (Fig 2b.). No interaction could be fitted between RIL and block due to the limited number of plants per RIL in each block. Because of the lack of interactions and the shared scaling exponent β among lines, the relative ranking of plant growth rates does not change between experimental blocks. The growth rates of the whole Kas x Tsu population had an approximately normal distribution and spanned a range of approximately 0.1 units (Fig. 2c).

We correlated SGR values with line means of glucosinolate concentration (Table 2). For correlations with compounds that are only produced by one of the two *GS-AOP* chemotypes, we only used the subset of lines producing the compound. In addition, we summed individual glucosinolate compounds into larger biosynthetic groups and correlated these with SGR values. We grouped compounds into total indolic and total aliphatic glucosinolates, and further split aliphatic glucosinolates according to the length of their carbon side-chain (3C, 4C or 8C). Despite the absence of a functional *4C*-allele in the Kas x Tsu population, we measured small concentrations of 4C glucosinolates. Several aliphatic glucosinolates showed significant positive correlations with growth. In contrast, the most abundant indolic glucosinolate, indolyl-3-methyl, and the total concentration of indolic glucosinolates were

negatively correlated with growth rate. As an alternative to correlations with individual compounds, we carried out a principal component analysis on the full set of glucosinolate compounds. However, this approach failed to reduce complexity of the data by a considerable amount, with eight out of fourteen principal components required to account for 90% of the variance, and was thus rejected (not shown).

Herbivory by Generalist and Specialist Caterpillars. The generalist caterpillar *Spodoptera littoralis* consumed 9.23 ± 4.7 mm^2 (mean ± 1 SD, 12.5 %) of leaf area within 24h, while the specialist *Plutella xylostella* consumed 10.66 ± 5.6 mm^2 (mean ± 1 SD, 15.39 %) in the same time. The percentage of rosette area removed by generalist and specialist caterpillars was positively correlated among lines (r = 0.641, n = 70, $P < 0.001$), i.e., there is some commonality in preference between the two herbivore species. However, despite this commonality there was a marginally significant interaction between herbivore identity and chemotype ($F_{1,593} = 3.31$, $P = 0.069$) on the amount of rosette area removed. *P. xylostella* tended to cause more damage to *ALK*-chemotypes than to *OHP*-chemotypes, while *S. littoralis* showed no such preference (Fig. 3). Of the 32 individual correlations, only two were significant (similar to the number expected by chance), with no obvious direction of the effects (Table 2). While part of this lack of significance can potentially be attributed to the limited number of lines used in the subset, it is in stark contrast to the results of Kliebenstein et al. (2002), who found strong negative correlations between glucosinolate compounds of L*er* x Cvi RILs and damage by the lepidopteran generalist herbivore *Trichoplusia ni*.

In an attempt to explain the discrepancies between the two studies, we compared our results with the published results of Kliebenstein et al. (2002, Fig. 4). Fig. 4 reveals the large difference in chemical profiles between the two RIL populations. In addition to the expected lack of 4C glucosinolates, the Kas x Tsu population also has reduced concentrations of 7-

carbon glucosinolates. The difference in the correlation patterns among glucosinolates, growth, and herbivore resistance could thus be due to the dramatically different chemotypes, in addition to the different species of generalist herbivores used.

Comparison of two RIL Populations. Plants of the Kas x Tsu RIL population grew considerably faster on average than plants from the L*er* x Cvi RIL population (Fig. 5a, t = 5.83, df = 1395, $P < 0.001$). This population-level difference can only be due to the *GS-ELONG* locus, as L*er* x Cvi lines have both *3C*- and *4C*-alleles, while all Kas x Tsu lines carry the *3C*-allele. In contrast, the *ALK* and *OHP* alleles at the *GS-AOP* locus are present in equal frequencies in both populations. *GS-ELONG* alone could account for some, but not all, of the difference between the two populations (Table 1). However, the best model included additive effects of both *GS-ELONG* and RIL population on all three model parameters. The *4C*-allele at the *GS-ELONG* locus had a significant negative effect on the rate parameter α (t = 2.74, df = 1392, $P = 0.006$); therefore, L*er* x Cvi lines accumulating 4C compounds were the slowest growing plants in the experiment (Fig. 5b).

DISCUSSION

Growth phenotyping of *Arapidopsis thaliana* based on automated, non-destructive plant imaging is currently receiving increased attention from plant scientists (e.g., Durham Brooks et al., 2010; Arvidsson et al., 2011). However, these studies rely on custom-built robotic systems that are affordable only to a small proportion of the scientific community. Here we present a method that requires nothing more than a standard camera and a software package, essentially producing the same results. This method allowed us to measure the growth rates of a complete RIL population allowing comparison of growth rates with other plant traits. SGR,

which is an estimate of relative growth rate unbiased by initial size, is expected to trade-off against costly traits such as defensive secondary metabolites. It was therefore somewhat surprising that the SGR values calculated from rosette areas in the Kas x Tsu RIL population showed few negative correlations with leaf glucosinolate content. This is in contrast to a previous study where we demonstrated strong negative correlations between biomass-based SGR values and leaf glucosinolate content using the L*er* x Cvi RIL population (Paul-Victor et al., 2010).

One possible explanation for the difference between these two studies lies with the technical details. Paul-Victor et al. (2010) measured SGR in L*er* x Cvi using destructive biomass measurements while in the present study we used non-destructive leaf area measurements. However, for a range of natural accessions, leaf area of *Arabidopsis* measured from photographs was highly correlated with plant biomass in the vegetative growth phase ($r = 0.92$, $P < 0.001$, Figure S1). As an additional test, we also re-analysed the biomass growth data presented in Paul-Victor et al. (2010) by restricting the dataset to the vegetative growth phase and fitting a power-law model to the data. Growth rates from the restricted L*er* x Cvi dataset were positively correlated with SGRs from the full dataset ($r = 0.74$, $P < 0.001$), and while the majority of correlations between compounds and herbivory scores were slightly less significant when using the reduced dataset, the direction of the effects remained the same. Methodological differences are thus unlikely to be solely responsible for the different results obtained for the two RIL populations.

In an attempt to explain the difference in growth/defence trade-offs between the two RIL populations, we compared the chemical profile of the Kas x Tsu population with the profile of the L*er* x Cvi population (Kliebenstein et al., 2002). The clearest difference between the two populations is the complete absence of aliphatic glucosinolates with a 4C side chain in the Kas x Tsu population due to the lack of a *4C*-allele at the *GS-ELONG* locus. In addition,

accumulation of aliphatic glucosinolates with a 7C side-chain is reduced in Kas x Tsu. In most genotypes of *A. thaliana*, and notably in the L*er* x Cvi population, all aliphatic glucosinolates are strongly positively correlated with each other (r = 0.36 – 0.71 for the main groups of aliphatic glucosinolates in L*er* x Cvi), making conclusions on the relative importance of individual groups of compounds difficult. The Kas x Tsu population is thus an elegant test of whether aliphatic 4C glucosinolates have particularly high growth costs. The standardised growth rate comparison of the two RIL populations revealed that Kas x Tsu lines did indeed grow faster on average than L*er* x Cvi lines. A substantial part of this difference was due to the presence of the *4C*-allele at *GS-ELONG,* present only in the L*er* x Cvi population. While costly to produce, these 4C glucosinolates may be particularly effective against generalist herbivores; for example Hansen et al. (2008) demonstrated that artificial up-regulation of an aliphatic 4C glucosinolate increased plant resistance towards the generalist herbivore *T. ni*. Additionally, the *GS-ELONG* locus has been linked to controlling plant development, circadian clock function and insect resistance in other studies (Kliebenstein et al., 2002; Kerwin et al., 2011).

Although the *4C*-allele at *GS-ELONG* confers a growth cost to the plant, there are other differences between Kas x Tsu and L*er* x Cvi that are not explained by chemotype. Growth differences between L*er* x Cvi lines carrying *3C*- and *4C*-alleles were most distinct in the early growth phase and decreased somewhat as plants approached the flowering phase. This supports the results of Züst et al. (2011) who found that the benefits of artificially disabled production of glucosinolates disappear as plants grow larger and switch to reproductive growth. It is therefore likely that costs of leaf-glucosinolates are most distinct in the early vegetative growth phase, when damage to leaves is most detrimental to the fitness of a plant.

In conclusion, the methods for non-destructive size measurement presented here allow growth phenotyping of the rosette stage of plants with high accuracy but comparably low costs. We employed these methods on a full RIL population consisting of 314 lines and were thus able demonstrate that aliphatic 4C glucosinolates are a prime example of a defensive trait according to plant defence theory: they are both an effective defence trait and costly to produce.

ACKNOWLEDGMENTS

This work was funded by the Forschungskredit of the University of Zürich (grant 57230502) to TZ, by Swiss National Science Foundation Grant 31-107531 to LAT, by US National Science Foundation grant IOS1021861, DBI0820580 and EF0735191 subaward Y551863 to DJK. We thank Arturo Bolaños and Prof. Ueli Grossniklaus for providing us with seeds of the L*er* x Cvi RIL population and Nathalie Séchaud and Sarah Pellkofer for help with the herbivory trials. Roland Reist of Syngenta Crop Protection Münchwilen AG kindly provided us with eggs of the two caterpillar species.

REFERENCES

Alonso-Blanco C, Peeters AJM, Koornneef M, Lister C, Dean C, van den Bosch N, Pot J, Kuiper MTR. 1998. Development of an Aflp Based Linkage Map of Ler, Col and Cvi *Arabidopsis Thaliana* Ecotypes and Construction of a Ler/Cvi Recombinant Inbred Line Population. *Plant Journal* **14**(2): 259-271.

Arvidsson S, Pérez-Rodríguez P, Mueller-Roeber B. 2011. A Growth Phenotyping Pipeline for *Arabidopsis Thaliana* Integrating Image Analysis and Rosette Area Modeling for Robust Quantification of Genotype Effects. *New Phytologist* **191**(3): 895-907.

Bazzaz FA, Chiariello NR, Coley PD, Pitelka LF. 1987. Allocating Resources to Reproduction and Defense. *Bioscience* **37**(1): 58-67.

Bidart-Bouzat MG, Kliebenstein DJ. 2008. Differential Levels of Insect Herbivory in the Field Associated with Genotypic Variation in Glucosinolates in *Arabidopsis Thaliana*. *Journal of Chemical Ecology* **34**(8): 1026-1037.

Bolker BM. 2008. *Ecological Models and Data in R*. Princeton, New Jersey: Princeton University Press.

Bones AM, Rossiter JT. 1996. The Myrosinase-Glucosinolate System, Its Organisation and Biochemistry. *Physiologia Plantarum* **97**(1): 194-208.

Coley PD, Bryant JP, Chapin FS. 1985. Resource Availability and Plant Antiherbivore Defense. *Science* **230**(4728): 895-899.

Durham Brooks TL, Miller ND, Spalding EP. 2010. Plasticity of *Arabidopsis* Root Gravitropism Throughout a Multidimensional Condition Space Quantified by Automated Image Analysis. *Plant Physiology* **152**(1): 206-216.

Enquist BJ, Brown JH, West GB. 1998. Allometric Scaling of Plant Energetics and Population Density. *Nature* **395**(6698): 163-165.

Enquist BJ, West GB, Charnov EL, Brown JH. 1999. Allometric Scaling of Production and Life-History Variation in Vascular Plants. *Nature* **401**(6756): 907-911.

Evans GC. 1972. *The Quantitative Analysis of Plant Growth*. Oxford: Blackwell Scientific Publications.

Fakheran S, Paul-Victor C, Heichinger C, Schmid B, Grossniklaus U, Turnbull LA. 2010. Adaptation and Extinction in Experimentally Fragmented Landscapes. *Proceedings of the National Academy of Sciences of the United States of America* **107**(44): 19120-19125.

Hansen BG, Kerwin RE, Ober JA, Lambrix VM, Mitchell-Olds T, Gershenzon J, Halkier BA, Kliebenstein DJ. 2008. A Novel 2-Oxoacid-Dependent Dioxygenase Involved in the Formation of the Goiterogenic 2-Hydroxybut-3-Enyl Glucosinolate and Generalist Insect Resistance in Arabidopsis. *Plant Physiology* **148**(4): 2096-2108.

Herms DA, Mattson WJ. 1992. The Dilemma of Plants: To Grow or Defend. *Quarterly Review of Biology* **67**(3): 283-335.

Ingestad T, Agren GI. 1992. Theories and Methods on Plant Nutrition and Growth. *Physiologia Plantarum* **84**(1): 177-184.

Kerwin RE, Jimenez-Gomez JM, Fulop D, Harmer SL, Maloof JN, Kliebenstein DJ. 2011. Network Quantitative Trait Loci Mapping of Circadian Clock Outputs Identifies Metabolic Pathway-to-Clock Linkages in *Arabidopsis*. *Plant Cell* **23**(2): 471-485.

Kliebenstein DJ, D'Auria JC, Behere AS, Kim JH, Gunderson KL, Breen JN, Lee G, Gershenzon J, Last RL, Jander G. 2007. Characterization of Seed-Specific Benzoyloxyglucosinolate Mutations in *Arabidopsis Thaliana*. *Plant Journal* **51**(6): 1062-1076.

Kliebenstein DJ, Gershenzon J, Mitchell-Olds T. 2001a. Comparative Quantitative Trait Loci Mapping of Aliphatic, Indolic and Benzylic Glucosinolate Production in *Arabidopsis Thaliana* Leaves and Seeds. *Genetics* **159**(1): 359-370.

Kliebenstein DJ, Kroymann J, Brown P, Figuth A, Pedersen D, Gershenzon J, Mitchell-Olds T. 2001b. Genetic Control of Natural Variation in *Arabidopsis* Glucosinolate Accumulation. *Plant Physiology* **126**(2): 811-825.

Kliebenstein DJ, Pedersen D, Barker B, Mitchell-Olds T. 2002. Comparative Analysis of Quantitative Trait Loci Controlling Glucosinolates, Myrosinase and Insect Resistance in *Arabidopsis Thaliana*. *Genetics* **161**(1): 325-332.

Lambrix V, Reichelt M, Mitchell-Olds T, Kliebenstein DJ, Gershenzon J. 2001. The Arabidopsis Epithiospecifier Protein Promotes the Hydrolysis of Glucosinolates to Nitriles and Influences *Trichoplusia Ni* Herbivory. *Plant Cell* **13**(12): 2793-2807.

Maranon T, Grubb PJ. 1993. Physiological-Basis and Ecological Significance of the Seed Size and Relative Growth-Rate Relationship in Mediterranean Annuals. *Functional Ecology* **7**(5): 591-599.

Mauricio R. 1998. Costs of Resistance to Natural Enemies in Field Populations of the Annual Plant *Arabidopsis Thaliana*. *American Naturalist* **151**(1): 20-28.

McKay JK, Richards JH, Nemali KS, Sen S, Mitchell-Olds T, Boles S, Stahl EA, Wayne T, Juenger TE. 2008. Genetics of Drought Adaptation in *Arabidopsis Thaliana* Ii. Qtl Analysis of a New Mapping Population, Kas-1 X Tsu-1. *Evolution* **62**(12): 3014-3026.

Mithen R, Clarke J, Lister C, Dean C. 1995. Genetics of Aliphatic Glucosinolates. Iii. Side-Chain Structure of Aliphatic Glucosinolates in *Arabidopsis Thaliana*. *Heredity* **74**: 210-215.

Paine CET, Marthews TR, Vogt DR, Purves D, Rees M, Hector A, Turnbull LA. in press. How to Fit Nonlinear Plant Growth Models and Calculate Growth Rates: An Update for Ecologists. *Methods in Ecology and Evolution*.

Paul-Victor C, Züst T, Rees M, Kliebenstein DJ, Turnbull LA. 2010. A New Method for Measuring Relative Growth Rate Can Uncover the Costs of Defensive Compounds in *Arabidopsis Thaliana*. *New Phytologist* **187**(4): 1102-1111.

Philipson CD, Saner P, Marthews TR, Nilus R, Reynolds G, Turnbull LA, Hector A. in press. Light-Based Regeneration Niches: Evidence from 21 Dipterocarp Species Using Size-Specific Rgrs. *Biotropica*.

R.Development.Core.Team 2010. R: A Language and Environment for Statistical Computing.In. Vienna, Austria: R Foundation for Statistical Computing. R Development Core Team.

Rasband WS 1997-2009. Imagej.In. U. S. National Institutes of Health, Bethesda, Maryland, USA.

Reichelt M, Brown PD, Schneider B, Oldham NJ, Stauber E, Tokuhisa J, Kliebenstein DJ, Mitchell-Olds T, Gershenzon J. 2002. Benzoic Acid Glucosinolate Esters and Other Glucosinolates from *Arabidopsis Thaliana*. *Phytochemistry* **59**(6): 663-671.

Rose KE, Atkinson RL, Turnbull LA, Rees M. 2009. The Costs and Benefits of Fast Living. *Ecology Letters* **12**: 1379-1384.

Thompson JN. 1999. Specific Hypotheses on the Geographic Mosaic of Coevolution. *American Naturalist* **153**: S1-S14.

Züst T, Joseph B, Shimizu KK, Kliebenstein DJ, Turnbull LA. 2011. Using Knockout Mutants to Reveal the Costs of Defensive Traits. *Proceedings of the Royal Society B: Biological Sciences* **278**: 2598-2603.

TABLES

TABLE 1. Comparison of models for the power-law regression model of plant growth for the growth rate comparison of two RIL populations. The first line is the basic model where all three model parameters (M_0, the growth coefficient α, and the scaling exponent β) are estimated for each population separately. Subsequent models test the effect of *GS-AOP* and *GS-ELONG* alleles on growth, and finally combine these with population effects if supported by the evaluation of F-tests. Stars (*) represent significant terms based on F-tests. The best model based on AIC value is marked in boldface.

starting size M_0	growth coefficient α	scaling exponent β	AIC
Population***	Population***	Population***	11980.92
AOP***	AOP***	AOP	14393.90
ELONG***	ELONG***	ELONG*	12031.25
AOP×ELONG***	AOP×ELONG***	AOP×ELONG	11990.17
Population***+AOP*	Population***+AOP	Population***+AOP	11986.72
Population*+ELONG****	**Population***+ELONG*****	**Population***+ELONG***	**11972.91**

TABLE 2. Correlations (Pearson's product moment) between concentration of glucosinolate compounds in leaves of *Arabidopsis thaliana* plants and the growth rate, generalist (*S. littoralis*) and specialist (*P. xylostella*) herbivory scores of the same genetic lines. Growth rate – glucosinolate correlations are based on the full 341 lines of the Kas x Tsu RIL population, while correlations with herbivory are based on the subset of 70 lines used for the herbivory experiment. Aliphatic 3C and 4C glucosinolate content is strongly affected by the *GS-AOP* locus, with some compounds only produced by one chemotype. '*GS-AOP* chemotype' indicates that correlations are only based on half the lines belonging to the respective group.

	GS-AOP chemotype	Size-standardised growth rate (SGR)		Herbivory by *S. littoralis*		Herbivory by *P. xylostella*	
3-hydroxypropyl	*OHP*	r = 0.104	p = 0.177	r = 0.082	p = 0.643	r = 0.113	p = 0.526
3-methylsulfinylpropyl	*OHP*	r = -0.006	p = 0.940	r = -0.133	p = 0.455	r = -0.043	p = 0.809
3-methylthiobutyl	*OHP*	r = -0.039	p = 0.614	r = -0.076	p = 0.669	r = -0.097	p = 0.584
Allyl	*ALK*	r = -0.004	p = 0.961	r = 0.062	p = 0.729	r = 0.081	p = 0.647
Total 3-carbon glucosinolates		r = 0.060	p = 0.270	r = 0.046	p = 0.703	r = 0.087	p = 0.474
4-hydroxybutyl	*OHP*	**r = 0.157**	**p = 0.041**	r = 0.254	p = 0.147	r = 0.107	p = 0.548
4-methylthiobutyl	*OHP*	r = 0.033	p = 0.670	r = 0.345	**p = 0.046**	r = 0.061	p = 0.731
Butenyl	*ALK*	r = 0.093	p = 0.230	r = 0.081	p = 0.648	r = 0.051	p = 0.776
Total 4-carbon glucosinolates		**r = 0.154**	**p = 0.004**	r = 0.164	p = 0.174	r = 0.098	p = 0.420
7-methylsulfinylheptyl		**r = 0.116**	**p = 0.031**	r = 0.203	p = 0.092	r = -0.008	p = 0.951
8-methylsulfinyloctyl		**r = 0.106**	**p = 0.050**	r = 0.204	p = 0.090	r = 0.141	p = 0.246
8-methylthiooctyl		r = 0.003	p = 0.953	r = 0.120	p = 0.323	r = 0.067	p = 0.581
Total 8-carbon glucosinolates		r = 0.086	p = 0.112	r = 0.208	p = 0.083	r = 0.139	p = 0.253
Total aliphatic glucosinolates		r = 0.071	p = 0.187	r = 0.086	p = 0.477	r = 0.103	p = 0.398
Indolyl-3-methyl		**r = -0.236**	**p < 0.001**	r =-0.105	p = 0.389	r = -0.250	**p = 0.037**
Total indolic glucosinolates		**r = -0.139**	**p = 0.010**	r = 0.142	p = 0.240	r = -0.017	p = 0.890

Significant correlations are presented in bold type

FIGURES

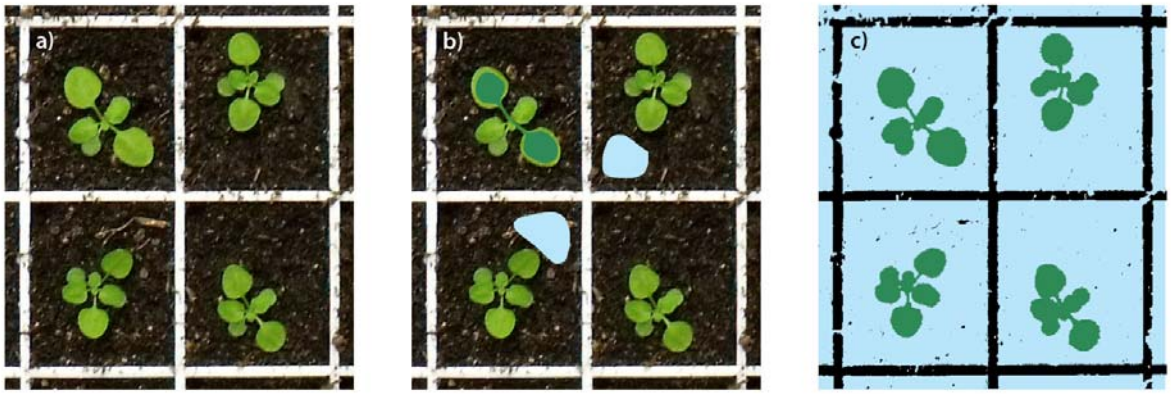


FIGURE 1. Section of a tray image before (a) and after (c) processing with ENVI. In (b), a typical example of a ‘training sample’ for maximum likelihood classification is shown, i.e., the region of interest which has to be drawn by the user to provide the software with a range of digital numbers for the targeted classes. After classification and quality control by the user, the position of each grid cell can be used to extract the leaf area of individual plants.

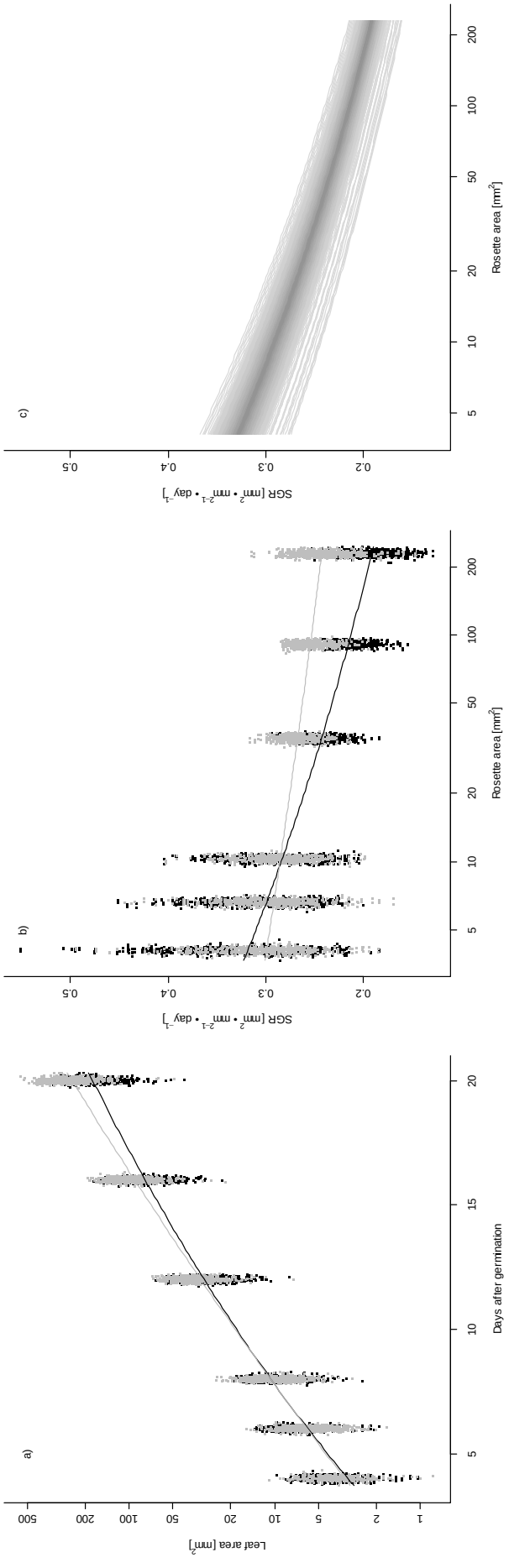


FIGURE 2. Power-law model fits for rosette area and SGR of the full RIL population. (a) Plot of rosette areas against time. Black and grey dots represent plants of the first and second experimental replicate, respectively. Fitted lines are the predicted overall means of each replicate. (b) Plot of size-standardised growth rate (SGR) against rosette size. Predicted rosette area represents the mean rosette area at the six measured days and is the reference size at which SGR is calculated. Black and grey dots are plant-specific SGR values of the first and second experimental replicate. Fitted lines are overall means of each replicate. (c) Plot of RIL-specific SGR against rosette size for all 356 lines, calculated for the first experimental replicate. Shades of gray are used to indicate the number of overlaid lines.

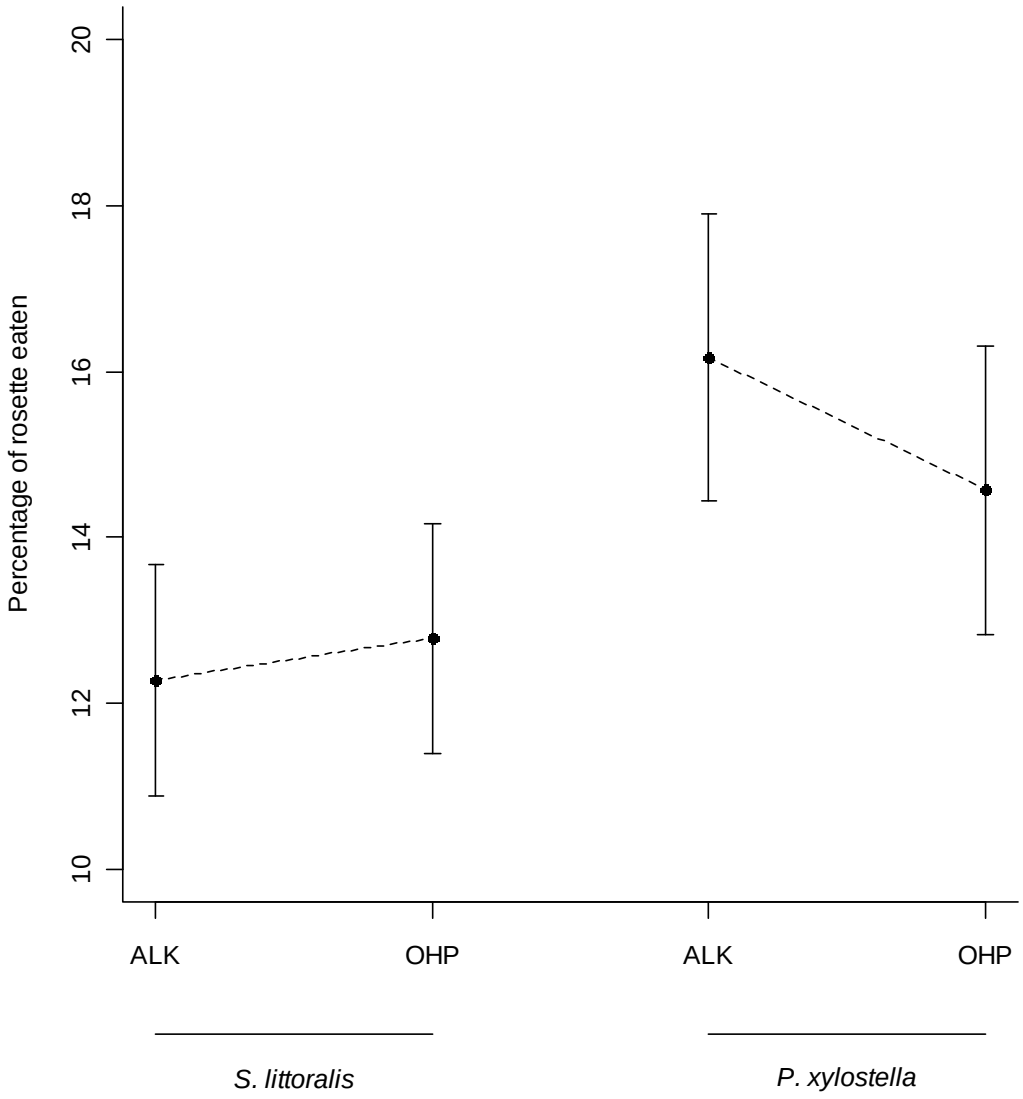


FIGURE 3. Mean percentage (with 95 % confidence interval) of removed rosette area after 24h of herbivory by the generalist *S. littoralis* and the specialist *P. xylostella* on a subset of 70 lines. Herbivory scores are presented for each chemotype separately.

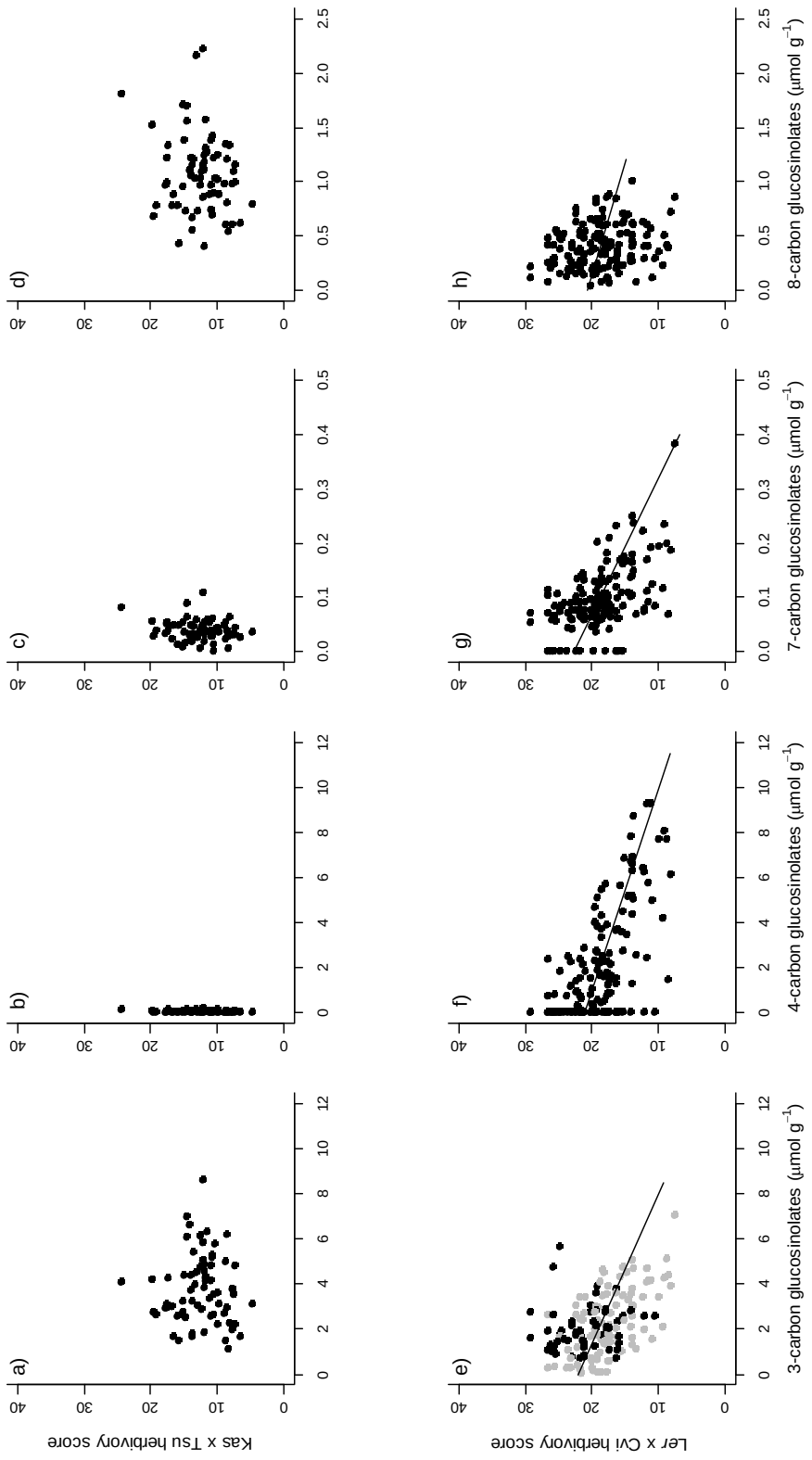


FIGURE 4. Percent leaf area removed plotted against concentrations of the four major groups of aliphatic glucosinolates. (a-d) Plots for a subset of 70 lines of the Kas x Tsu RIL population with *S. littoralis* as generalist herbivore. (e-f) For comparison, data from Kliebenstein et al. (2002) is plotted, showing herbivory on all 153 lines of the L*er* x Cvi RIL population with *T. ni* as generalist herbivore. Significant relations are indicated by solid lines. 3- to 8-carbon refers to the length of the carbon side-chain in aliphatic glucosinolates. In panel e), 4C producing lines are coloured in grey.

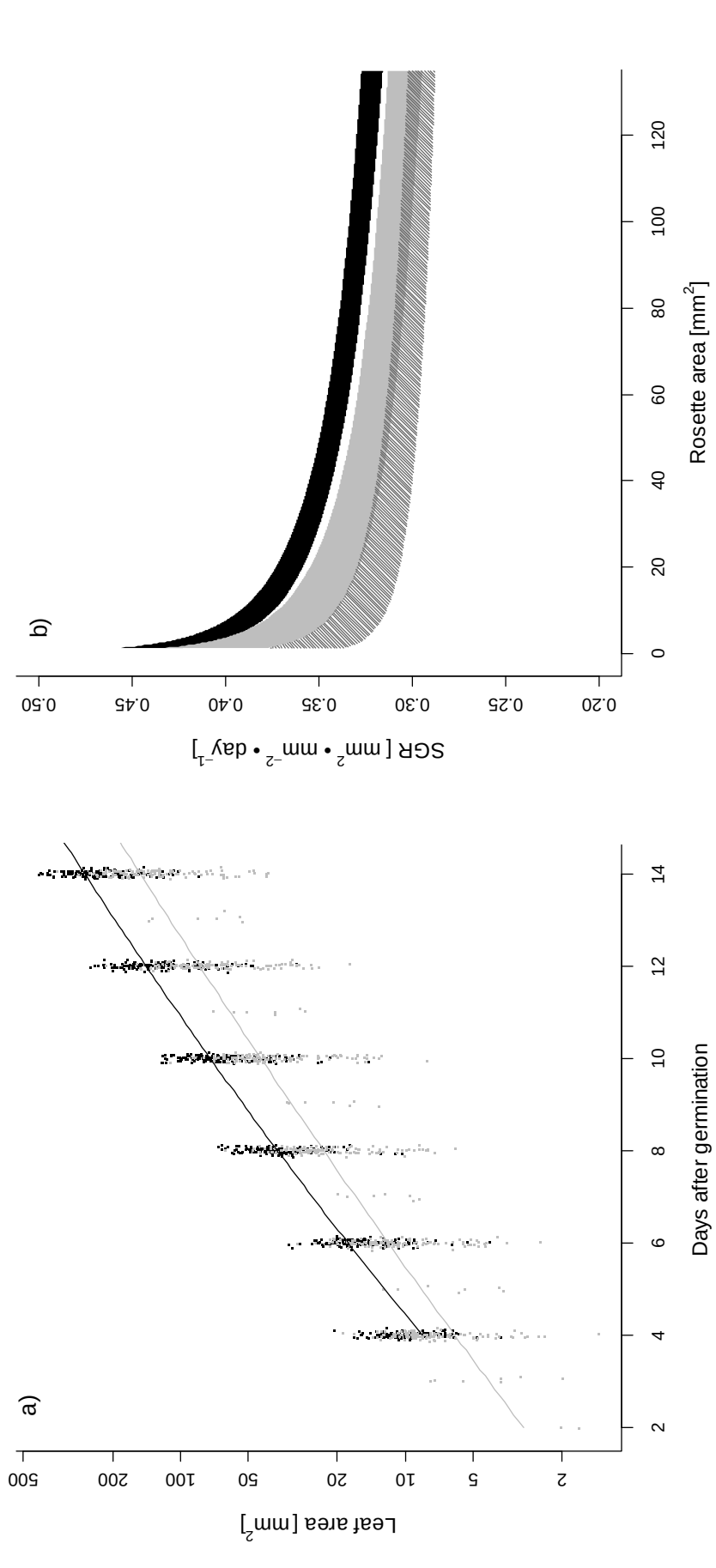


FIGURE 5. (a) Power-law model fit of rosette area against age (corrected for germination day) for a subset of 30 lines of the L*er* x Cvi RIL population (grey) and of 30 lines of the Kas x Tsu RIL population (black), grown as individual plants in the same experiment. Fitted lines are the predicted population means. (b) Plot of the size-standardised growth rate of the Kas x Tsu RIL population (black) and of the *3C* (grey) and *4C* (hatched) lines of the L*er* x Cvi population. Areas represent the 95 % prediction intervals around a population mean. SGRs are calculated for increasing rosette sizes, up to the mean rosette size of L*er* x Cvi lines at the last measured day to prevent prediction beyond the range of the data.

SUPPORTING INFORMATION

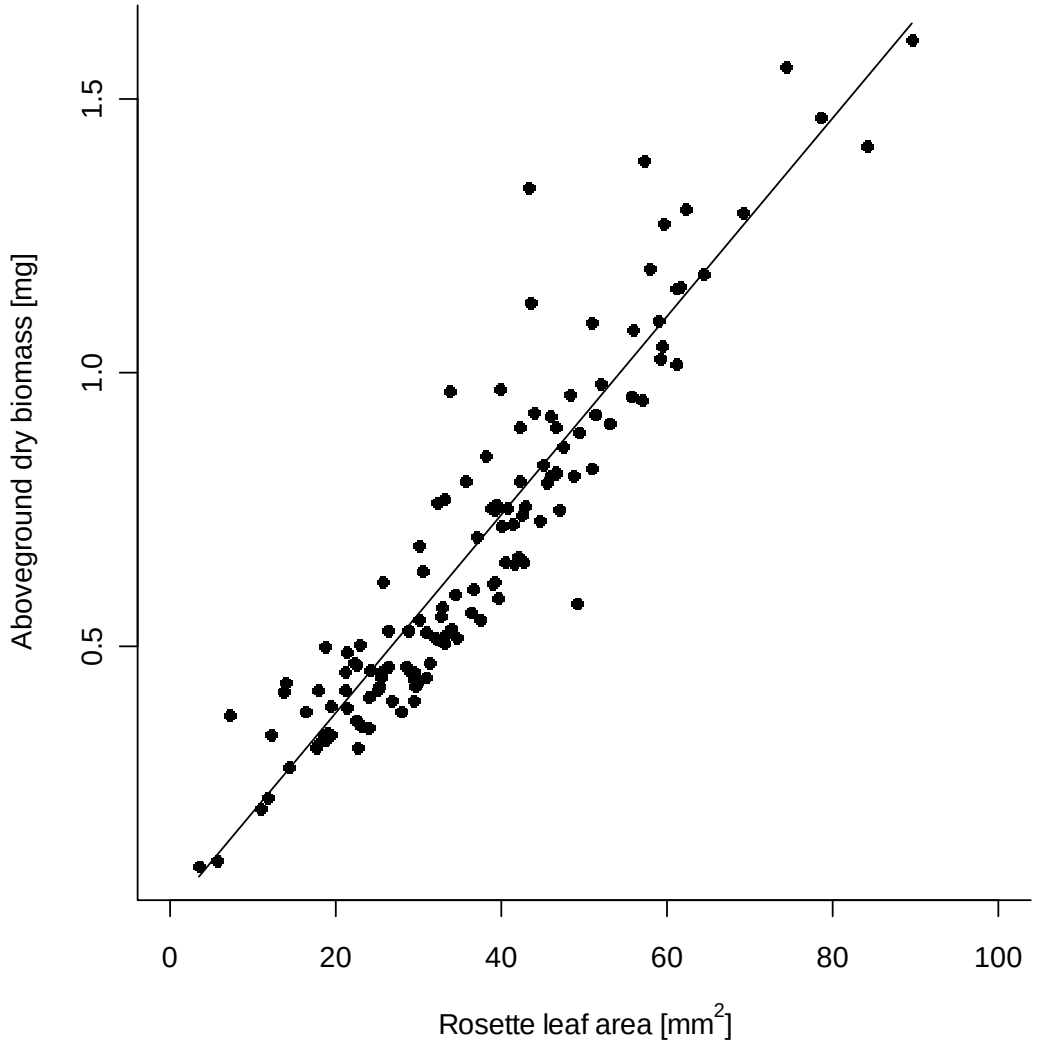


FIGURE S1. Plot illustrating the close linkage of leaf area measured from photographs and plant biomass. The plants displayed here are a selection of natural *Arabidopsis* accessions, measured 9 or 11 days after sowing. Plants were grown in 96-well planting trays and photographed immediately before the harvest.

Natural Enemies Drive Geographic Variation in Plant Defenses

Tobias Züst[1*†], Christian Heichinger[2], Ueli Grossniklaus[2], Richard Harrington[3], Daniel J. Kliebenstein[4,5] and Lindsay A. Turnbull[1]

Plants defend themselves against attack by natural enemies and these defenses vary widely across populations. However, whether communities of natural enemies are a sufficiently potent force to maintain polymorphisms in defensive traits is largely unknown. Here, we exploit the genetic resources of *Arabidopsis thaliana*, coupled with 39 years of field data on aphid abundance to: (i) demonstrate that geographic patterns in a polymorphic defense locus (*GS-ELONG*) are strongly correlated with changes in the relative abundance of two specialist aphids; and (ii) demonstrate differential selection by the two aphids on *GS-ELONG*, using a multi-generation selection experiment. We thereby show a causal link between variation in abundance of the two specialist aphids and the geographic pattern at *GS-ELONG*, which highlights the potency of natural enemies as selective forces.

Intraspecific genetic variation is essential in enabling species to respond rapidly to evolutionary challenges such as changing environmental conditions (*1*) or the emergence of novel pests and pathogens (*2*). This diversity often reflects the balance between the strength of local selection and the current and historical levels of population substructure and gene flow (*3, 4*). Geographic analyses of genetic variation in several plant species have revealed clear genetic signals of local adaptation (*5*), caused by differences in the selective regime among locations. These analyses are further supported by reciprocal transplant experiments, in which home genotypes generally outperform those transplanted from other populations (*6, 7*). While the drivers of local adaptation often remain unidentified, there is evidence that climate and soil can exert strong local selective pressures and play important roles in shaping large-scale genetic patterns (*8, 9*).

In contrast to the clear role of abiotic factors, there is little direct evidence that biotic forces, such as herbivory or competition, can lead to the maintenance of genetic variation across large geographic scales, despite the exceptional levels of polymorphism associated with genes involved in defense (*10, 11*). In theory, interactions between organisms and their natural enemies can lead to differences in the local selective regime because of geographic variation in the abundance or species composition of the enemy community (*3*). This spatial variation can affect defense if it is costly; e.g., if the average level of herbivory varies across populations, defended genotypes might dominate in heavily attacked populations while undefended genotypes would prosper when enemies are absent or rare (*12*). Another less studied effect is how defense might vary if plants are attacked by diverse collections of herbivore species that differ in feeding style and specialization. This could lead to higher levels of polymorphism in

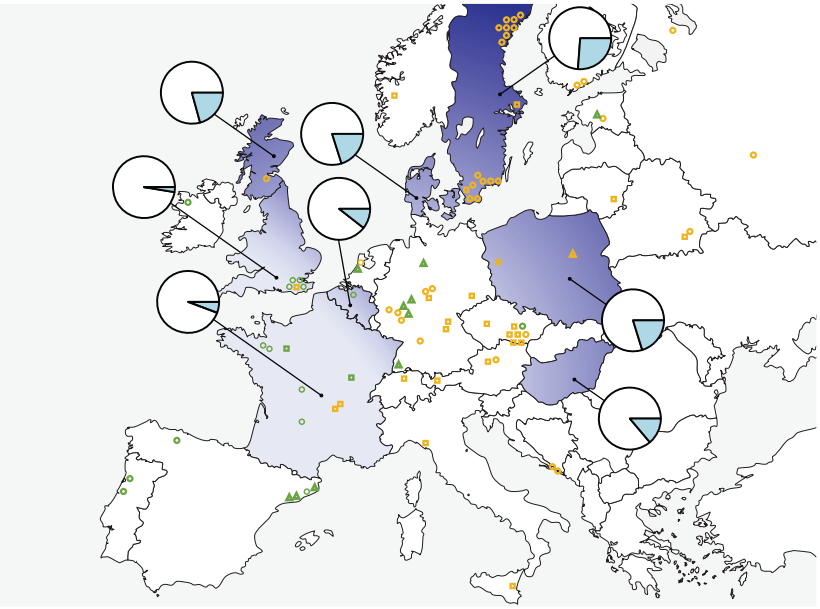


Fig. 1. Location of European *A. thaliana* accessions with known chemical profile. Symbol color indicates the *GS-ELONG* chemotype (orange: *3C*; green: *4C*) and symbol shape indicates the *GS-AOP* chemotype (square: *ALK*; circle: *OH*; triangle: *NULL*). For GS-ELONG the probability of finding *3C* populations increases strongly with longitude (binomial glm: $t = 5.11$, df = 85, $p < 0.001$) and weakly with latitude ($t = 1.75$, df = 85, $p = 0.084$). Countries with available aphid data are colored in blue. The shade of blue corresponds to the relative frequency of *L. erysimi* based on model predictions from a binomial GLM using data from 61 aphid suction traps. The relative frequency of *L. erysimi* increases strongly with longitude ($t = 5.03$, $p < 0.001$) and weakly with latitude ($t = 1.89$, $p = 0.060$). Piecharts indicate the observed average relative abundance of *B. brassicae* (white) and *L. erysimi* (blue) in each country.

[1] Institute of Evolutionary Biology and Environmental Studies and Zürich-Basel Plant Science Center, University of Zürich, Zürich, CH-8057, Switzerland. [2] Institute of Plant Biology and Zürich-Basel Plant Science Center, University of Zürich, Zürich, CH-8008, Switzerland. [3] Rothamsted Insect Survey, AgroEcology Department, Rothamsted Research, Harpenden AL5 2JQ, UK. [4] Department of Plant Sciences, University of California at Davis, Davis, California 95616, USA. [5] DynaMo Center of Excellence, Department of Plant Biology and Biotechnology, University of Copenhagen, DK-1871, Denmark.

*To whom correspondence should be addressed. E-mail: tobias.zuest@cornell.edu.

†Present address: Department of Ecology and Evolutionary Biology, Cornell University, Ithaca, NY 14853, USA.

defense genes due to selection for specific defensive profiles matched to the predominant local herbivore or herbivore community (e.g., *13*). However, there is no direct evidence that variation in local herbivore communities represents a sufficiently strong selective pressure to favor specific defensive traits and maintain polymorphisms in defense-related genes.

The unparalleled genetic and molecular resources available for the model plant *Arabidopsis thaliana* make this species an ideal candidate to study the process of local adaptation to herbivores. The primary defensive trait in *A. thaliana* is a series of indolic and aliphatic glucosinolates, which are secondary plant metabolites with anti-herbivore properties (*14*). The accumulation and structure of aliphatic glucosinolates is mechanistically determined by alleles at the *GS-ELONG* locus that regulate the carbon side-chain elongation (*3C* or *4C*) (*15*) and by alleles at the *GS-AOP* locus that modify the functional group of the biologically active glucosinolate side-chain (*ALK*, *OH*, or *NULL*). The combination of these alleles yields six distinct chemotypes present in natural populations in varying proportions (*16*). Both individual glucosinolate compounds and full chemical profiles affect the susceptibility of a plant to specific herbivores (*17, 18*); hence the aliphatic chemotype is likely under differential, qualitative selection by herbivores. In contrast, accumulation of the main indolic glucosinolates in *A. thaliana* is highly plastic and modulated by a large number of small-effect genetic loci, which are therefore less likely to show clear signatures of selection (*19*).

We mapped geographic variation in the abundance of the six chemotypes within Europe from a set of 96 accessions (75 European) (*20*) with known chemical profiles (*16*) (Fig. 1). There was no apparent pattern in the distribution of the *GS-AOP* chemotypes, but for *GS-ELONG* the frequency of 3C to 4C chemotypes increases with both latitude and longitude (Figs 1, S1). If this pattern results from geographical variation in herbivore feeding pressure, we would expect it to be closely matched by variation in herbivore abundance patterns. While *A. thaliana* is attacked by a range of invertebrate herbivores, many of which preferentially feed on specific chemotypes (*17*), we hypothesized that the aphid species *Brevicoryne brassicae* and *Lipaphis erysimi* are likely drivers of these patterns as they are both abundant, mobile Brassicaceae specialists, yet differentially sensitive to environmental conditions (*21*). Fluctuations in aphid populations have been monitored since 1964 through the EXAMINE network (http://www.rothamsted.ac.uk/examine/) using suction traps which operate throughout the aphid flight season (*22*). We retrieved data on the two aphid species from 61 traps in eight European countries. These data revealed that the abundance of *L. erysimi* is usually lower than *B. brassicae*, but that the geographic pattern in the relative abundances of *L. erysimi* and *B. brassicae* closely mirrors the pattern at GS-ELONG (Figs 1 and S1). Variation in the relative abundance of these two specialist aphids could therefore underlie variation in the predominant *GS-ELONG* chemotype found in natural populations.

Since causal inferences are impossible from such correlative data, we tested the causality of aphid selection on *GS-ELONG*, carrying out a multi-generational selection experiment on populations of *A. thaliana* (*22*). We assembled 30 replicate populations from equal numbers of seeds from each of 27 natural accessions, including 6 of the 75 European accessions mapped above. Accessions were chosen to maximize variation in defense traits while including all six glucosinolate chemotypes in a range of genetic backgrounds (Table S1, Fig. S2). Over five generations, we consistently exposed populations to replicate (n = 6) treatments of a single specialist aphid species: either *B. brassicae* or *L. erysimi*; a single generalist aphid, *Myzus persicae*; a mixture of all three aphid species; and a no-aphid treatment. The generalist aphid was included as a negative control, since *M. persicae* is unresponsive to aliphatic glucosinolates (*23*) and we therefore would not expect it to exert directional selection on plant chemotype. The no-aphid treatment served as a control for other selective forces that were likely to affect the outcome of the experiment, such as intraspecific competition among accessions. Seeds were collected at the end of each generation with no mixing among populations and a subset was used to establish the next generation at a constant density. After five generations of repeated herbivore treatments we sampled 24 individuals from each population in generation 5 (144 individuals per treatment) and determined their genotype. To have a marker for changes in genotypic composition through time, we also measured leaf trichome density, a trait under strong genetic control (Fig. S3), on a representative sample of plants in all generations.

Rapid adaptation occurred in the selection experiment, as evidenced by a progressive reduction in the effects of aphid feeding on final plant biomass in each generation (Fig. 2A). In line with the expected severity of aphid feeding based on previously reported population growth rates (*21*), *L. erysimi* caused the strongest reduction in plant biomass, while *M. persicae* was intermediate and *B. brassicae* had the least effect. The mixture treatment caused a similar reduction to *L. erysimi* alone, probably because aphid mixtures were dominated by this fast-growing aphid species. With each generation, trichome density decreased in the no-aphid treatment, while it remained at significantly higher levels in all aphid treatments (Fig. 2B). Adaptation to herbivore feeding was accompanied by considerable changes in the genotypic composition of populations, including the complete loss of nine genotypes (Fig. 3). There was a non-specific aphid effect on total indolic glucosinolates (lme: $F_{1,28}$ = 10.66, p =0.003), with plants in the no-aphid treatment producing on average 0.98 (± 0.03, SEM) µmol g^{-1}, and plants in aphid treatments producing 0.87 (± 0.03, SEM) µmol g^{-1}. In contrast, the different aphid treatments had a dramatic impact on the dominant aliphatic chemotypes within experimental populations. Significantly, the relative proportions of *3C* and *4C* chemotypes differed strongly among aphid treatments (Figs. 3, S4). After selection, populations of

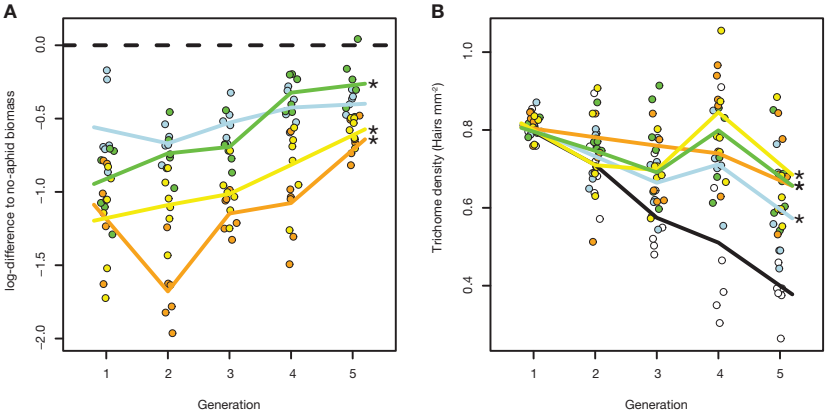


Fig. 2. A) Change in the negative impact of aphid treatments on final plant biomass over five generations, displayed as the log-difference to final plant biomass in the no-aphid treatment: *B. brassicae* (light blue); *M. persicae* (light green); *L. erysimi* (orange); and aphid mixture (yellow). Stars denote significantly less damage after five generations of selection (Table S2). B) Mean number of trichomes on the fourth leaf of 50 plants per population. Stars denote significant difference from the no-aphid treatment (black line) after five generations of selection.

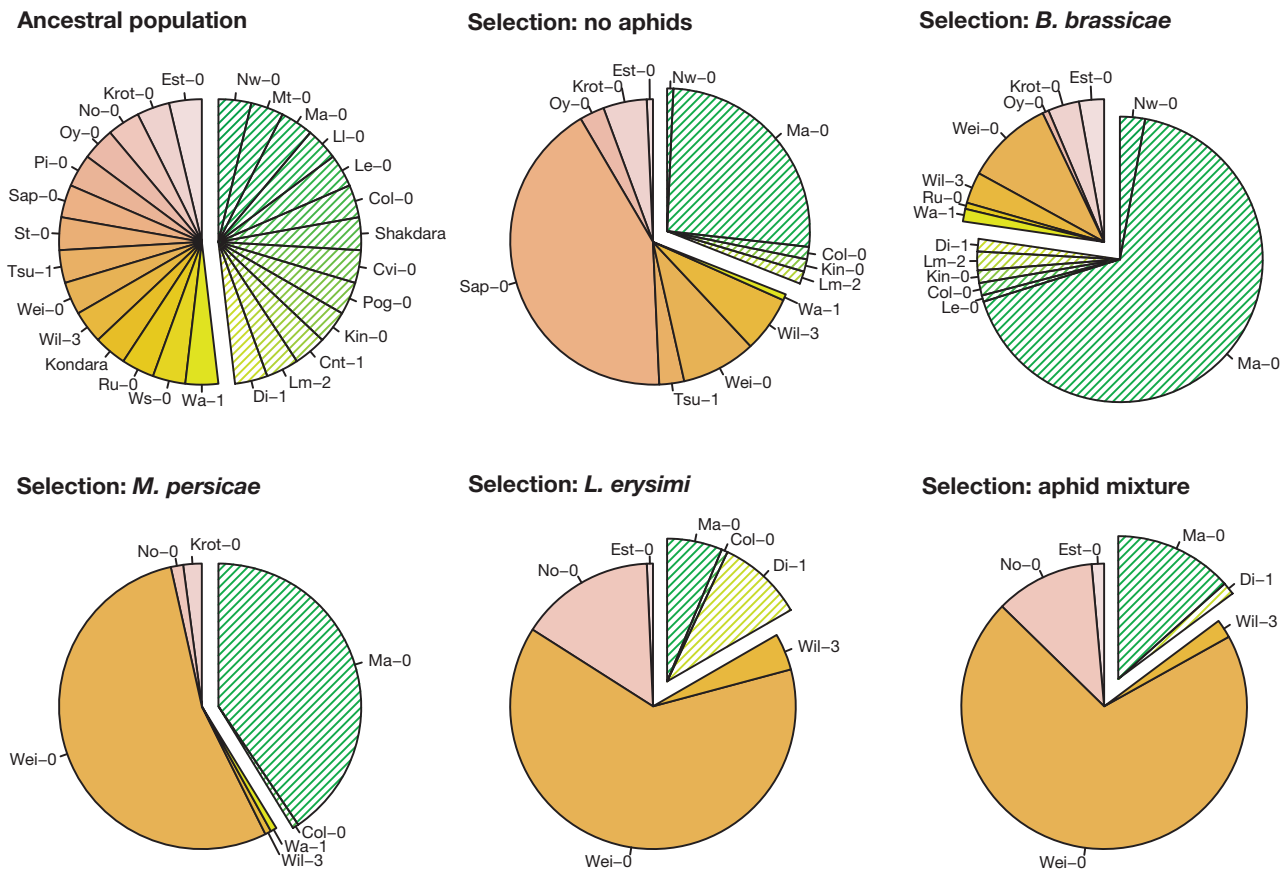


Fig. 3. Change in the composition of *A. thaliana* accessions, from equal proportions of 27 genotypes in the ancestral population to treatment-specific compositions after five generations of selection. Each chart gives mean genotype frequencies based on $n = 6$ replicate populations. *3C* chemotypes are indicated by solid, orange colors, while *4C* chemotypes are indicated by hatched, green colors.

the no-aphid treatment consisted of approximately two thirds *3C* and one third *4C* chemotypes. Specialist aphids selected for different chemotypes at *GS-ELONG*: the *4C* chemotypes strongly dominated in *B. brassicae* treatments (binomial glm, $t = 3.08$, df = 25, p = 0.002) and the *3C* chemotypes strongly dominated in both *L. erysimi* ($t = 2.01$, df = 25, p = 0.045) and the aphid mixture treatments ($t = 2.21$, df = 25, p = 0.027). The relative proportions of *3C* to *4C* chemotypes in populations exposed to the generalist aphid *M. persicae* did not differ from the no-aphid treatment ($t = 0.18$, df = 25, p = 0.858). Despite this similarity, the identity of the successful genotypes differed among treatments, with accession Sap-0 accounting for a large fraction of plants in the no-aphid treatment but being absent from all other treatments (Fig. 3). The genotypic composition of plant populations with *L. erysimi* and aphid mixtures was near-identical, confirming that *L. erysimi* dominated the mixture treatments, and suggesting that in co-founded populations, *L. erysimi* is the most important selective force. Most successful genotypes either had a *3C-OH* or a *4C-NULL* chemotype, and we found no individuals belonging to either alkenyl chemotype (*3C-ALK* or *4C-ALK*) in any treatment. Alkenyl chemotypes were common in generation 1 of the selection experiment (Fig. S2), and simulations of random sampling on the basis of observed population sizes reveal that their loss cannot be due to drift alone (Fig. S5) but rather was a consequence of selection.

To identify potential causes for the loss of particular genotypes, we measured size-standardized growth rate (SGR) as a measure of fitness, together with total aliphatic glucosinolate content and trichome density in a separate experiment on all 27 ancestral accessions. Alkenyl chemotypes expressed the highest levels of glucosinolates and were among the slowest growing genotypes overall (Fig. S6A). Alkenyl glucosinolates are an effective defense against leaf-chewing herbivores such as caterpillars (*24*), but their efficiency against specialist aphids remains largely unknown, while they have little effect on *M. persicae* (*23*). The loss of the alkenyl chemotypes therefore probably resulted from selection against a costly defense trait that provided insufficient benefits in our experiment. This cost-benefit balance is also the most likely reason for the difference in dominant genotypes between the no-aphid treatment and the aphid treatments (Fig. 3). The dominant genotype in no-aphid populations, Sap-0, was completely absent from all aphid treatments, indicating low fitness in the presence of herbivores. The Sap-0 genotype had the lowest trichome density of all non-glabrous accessions, and as trichome production had a growth cost (Fig S6B), its success can explain the observed decrease in trichome density in the no-aphid populations over time (Fig. 2B). Compared to other chemotypes, Sap-0 also produces low levels of glucosinolates, an additional indication that in the absence of herbivores, undefended, fast-growing genotypes will prosper.

Despite known epistatic interactions between *GS-ELONG* and *GS-AOP* (*19*), our data suggest that aphid selection acts independently on the two loci. The magnitude and direction of selection exerted by the two specialist aphids on *GS-ELONG* in our experiment suggests a causal link between the observed cline in *GS-ELONG* across Europe and the changes in the relative abundance of the same aphids. Although *B. brassicae* is numerically

dominant across most of Europe, the faster-growing *L. erysimi* can inflict greater damage on plants and quickly dominates populations which are co-founded by both aphid species, thus even a modest change in relative abundance could cause the loss of *C3* populations. All plants in the selection experiment experienced strong intraspecific competition, and since growth rate is a good predictor of competitive ability (*25*) it is unsurprising that fast-growing plant genotypes were generally selected, while the slowest-growing alkenyl chemotypes were lost. Alkenyl chemotypes are, however, very common in natural populations, and could be maintained by other herbivores, for instance leaf-chewing caterpillars (*24*).

Ecological theory has consistently emphasized the role of natural enemies in maintaining diversity both within and among species, but convincing empirical evidence has been lacking. Here we demonstrate that even functionally similar herbivores such as different species of aphid have the potential to select for specific chemotypes and drive large-scale geographic patterns in plant defense profiles. It therefore seems likely that natural herbivore communities with their greater variety of feeding styles and specializations play a major role in shaping and refining the plant defenses observed in natural communities.

References and Notes:

1. G. Bell, A. Gonzalez, Adaptation and evolutionary rescue in metapopulations experiencing environmental deterioration. *Science* 332, 1327-1330 (2011).
2. M. C. Fisher *et al.*, Emerging fungal threats to animal, plant and ecosystem health. *Nature* 484, 186-194 (2012).
3. B. Charlesworth, M. Nordborg, D. Charlesworth, The effects of local selection, balanced polymorphism and background selection on equilibrium patterns of genetic diversity in subdivided populations. *Genet. Res.* 70, 155-174 (1997).
4. M. Slatkin, Gene flow and the geographic structure of natural populations. *Science* 236, 787-792 (1987).
5. Y. B. Linhart, M. C. Grant, Evolutionary significance of local genetic differentiation in plants. *Annu. Rev. Ecol. Syst.* 27, 237-277 (1996).
6. A. Fournier-Level *et al.*, A map of local adaptation in *Arabidopsis thaliana*. *Science* 334, 86-89 (2011).
7. R. Leimu, M. Fischer, A meta-analysis of local adaptation in plants. *Plos One* 3 (2008).
8. N. Alvarez *et al.*, History or ecology? Substrate type as a major driver of spatial genetic structure in Alpine plants. *Ecol. Lett.* 12, 632-640 (2009).
9. A. M. Hancock *et al.*, Adaptation to climate across the *Arabidopsis thaliana* genome. *Science* 333, 83-86 (2011).
10. J. Bergelson, M. Kreitman, E. A. Stahl, D. C. Tian, Evolutionary dynamics of plant R-genes. *Science* 292, 2281-2285 (2001).
11. R. M. Clark *et al.*, Common sequence polymorphisms shaping genetic diversity in *Arabidopsis thaliana*. *Science* 317, 338-342 (2007).
12. S. Y. Strauss, J. A. Rudgers, J. A. Lau, R. E. Irwin, Direct and ecological costs of resistance to herbivory. *Trends Ecol. Evol.* 17, 278-285 (2002).
13. M. T. J. Johnson, A. A. Agrawal, Plant genotype and environment interact to shape a diverse arthropod community on evening primrose (*Oenothera biennis*). *Ecology* 86, 874-885 (2005).
14. B. A. Halkier, J. Gershenzon, Biology and biochemistry of glucosinolates. *Annu. Rev. Plant Biol.* 57, 303-333 (2006).
15. J. Kroymann, S. Donnerhacke, D. Schnabelrauch, T. Mitchell-Olds, Evolutionary dynamics of an *Arabidopsis* insect resistance quantitative trait locus. *Proc. Natl. Acad. Sci. USA* 100, 14587-14592 (2003).
16. E. K. F. Chan, H. C. Rowe, J. A. Corwin, B. Joseph, D. J. Kliebenstein, Combining genome-wide association mapping and transcriptional networks to identify novel genes controlling glucosinolates in *Arabidopsis thaliana*. *PLoS Biol.* 9 (2011).
17. M. G. Bidart-Bouzat, D. J. Kliebenstein, Differential levels of insect herbivory in the field associated with genotypic variation in glucosinolates in *Arabidopsis thaliana*. *J. Chem. Ecol.* 34, 1026-1037 (2008).
18. E. L. Newton, J. M. Bullock, D. J. Hodgson, Glucosinolate polymorphism in wild cabbage (*Brassica oleracea*) influences the structure of herbivore communities. *Oecologia* 160, 63-76 (2009).
19. D. J. Kliebenstein, J. Gershenzon, T. Mitchell-Olds, Comparative quantitative trait loci mapping of aliphatic, indolic and benzylic glucosinolate production in *Arabidopsis thaliana* leaves and seeds. *Genetics* 159, 359-370 (2001).
20. M. Nordborg *et al.*, The pattern of polymorphism in *Arabidopsis thaliana*. *PLoS Biol.* 3, 1289-1299 (2005).
21. C. J. DeLoach, Rate of increase of populations of cabbage, green peach, and turnip aphids at constant temperatures. *Ann. Entomol. Soc. Am.* 67, 332-340 (1974).
22. Materials and methods are available as supplementary material on Science Online.
23. J. H. Kim, G. Jander, *Myzus persicae* (green peach aphid) feeding on Arabidopsis induces the formation of a deterrent indole glucosinolate. *Plant J.* 49, 1008-1019 (2007).
24. D. J. Kliebenstein, Secondary metabolites and plant/environment interactions: a view through *Arabidopsis thaliana* tinged glasses. *Plant Cell Environ.* 27, 675-684 (2004).
25. S. Fakheran *et al.*, Adaptation and extinction in experimentally fragmented landscapes. *Proc. Natl. Acad. Sci. USA* 107, 19120-19125 (2010).
26. R. Harrington *et al.*, in *Aphids in a new millennium* J. C. Simon, C. A. Dedryver, C. Rispe, M. Hullé, Eds. (Proceedings 6th International Aphid Symposium, Rennes, France, 2004) pp. 45-49.
27. L. R. Taylor, Synoptic dynamics, migration and the Rothamsted insect survey - presidential address to the British Ecological Society, December 1984. *J. Anim. Ecol.* 55, 1-38 (1986).
28. E. D. M. Macaulay, G. M. Tatchell, L. R. Taylor, The Rothamsted insect survey 12-metre suction trap. *B. Entomol. Res.* 78, 121-129 (1988).
29. R. Harrington *et al.*, Environmental change and the phenology of European aphids. *Glob. Change Biol.* 13, 1550-1564 (2007).
30. N. Cocu, R. Harrington, M. Hullé, M. D. A. Rounsevell, Spatial autocorrelation as a tool for identifying the geographical patterns of aphid annual abundance. *Agr. Forest Entomol.* 7, 31-43 (2005).
31. D. J. Kliebenstein *et al.*, Genetic control of natural variation in *Arabidopsis* glucosinolate accumulation. *Plant Physiol.* 126, 811-825 (2001).
32. J. Hagenblad *et al.*, Haplotype structure and phenotypic associations in the chromosomal regions surrounding two *Arabidopsis thaliana* flowering time loci. *Genetics* 168, 1627-1638 (2004).
33. S. Atwell *et al.*, Genome-wide association study of 107 phenotypes in *Arabidopsis thaliana* inbred lines. *Nature* 465, 627-631 (2010).
34. C. J. Bell, J. R. Ecker, Assignment of 30 microsatellite loci to the linkage map of *Arabidopsis*. *Genomics* 19, 137-144 (1994).
35. K. Edwards, C. Johnstone, C. Thompson, A simple and rapid method for the preparation of plant genomic DNA for PCR analysis ANALYSIS. *Nucleic Acids Res.* 19, 1349-1349 (1991).
36. R Development Core Team. R Foundation for Statistical Computing, Vienna, Austria (2010).
37. W. S. Rasband. U. S. National Institutes of Health, Bethesda, Maryland, USA (1997-2009).
38. T. Züst, B. Joseph, K. K. Shimizu, D. J. Kliebenstein, L. A. Turnbull, Using knockout mutants to reveal the costs of defensive traits. *Proc. R. Soc. B* 278, 2598-2603 (2011).
39. J. C. Pinheiro, D. M. Bates, *Mixed-Effects Models in S and S-Plus*, Statistics and Computing (Springer, London, 2000).

Acknowledgments: This work was supported by the Forschungskredit of the University of Zürich to TZ and by Swiss National Science Foundation Grant 31-107531 to LAT, by US National Science Foundation grant DBI0820580 to DJK, an Advanced Grant of the European Research Council to UG, and a special grant of the University of Zürich in the commemoration of Prof. Christine B. Müller. The Rothamsted Insect Survey is supported by a BBSRC National Capability Grant and the Lawes Agricultural Trust. We thank members of the EU EXAMINE Project (EVK2-1999-00151) for use of data and A. A. Agrawal, S. West, J. Levine, and two anonymous reviewers for comments that improved the manuscript. All data presented in this paper and R code of all analyses are provided in the Supporting Online Material.

Supplementary Materials:

www.sciencemag.org
Materials and Methods
Figs. S1-S7
Tables S1, S2, S3
File with R code
Datafiles s3-s8

Supplementary Materials:

Additional Author notes: TZ and LAT proposed the original idea, performed statistical analyses, and wrote the paper; CH, UG, RH and DJK contributed data or provided molecular and chemical analyses; all authors contributed to revisions; TZ created the figures.

Materials and Methods:

Analysis of geographic patterns. We extracted coordinates data on the set of 96 natural accessions from Nordborg et al. (*20*) and combined it with glucosinolate data from Chan et al. (*16*). We analyzed the spatial pattern in *GS-ELONG* for Europe using a generalized linear model with binomial error structure and longitude and latitude as explanatory variables. There was no support for an interaction term; hence only main effects were included in the model. Data on aphid abundance were extracted from the database of the 'EXAMINE' network (*26*). This project emerged from the Rothamsted Insect Survey (*27*), which uses a network of 12.2 m high suction traps (*28*) throughout the UK to monitor aphids for research and extension work. Daily data are available for a large number of aphid species back to 1964 depending on the site. Other countries have adopted the same design of trap and the data have been assembled within a single database (*26*), facilitating pan-European analysis (*29*). There are currently 46 traps operating in ten countries, each representative of aphids flying within a surrounding area with a radius of more than 200 km (*30*). Many more traps have been operated in the past but have now been closed. This paper draws on samples from a total of 61 traps with varying extent of data coverage. Data from any particular trap in a given year were only used if at least one individual of both *Brevicoryne brassicae* and *Lipaphis erysimi* each were recorded in that location, since lack of data can be caused both by true absence of an aphid or by failure to monitor or record this species. We analyzed the aphid data with a binomial generalized mixed effects model (binomial glmer) for proportion data by treating the

total numbers of *L. erysimi* and *B. brassicae* as successes and failures, respectively. Trap identity was included as a random effect to account for both repeated measures and differences in the range of annual data. Longitude and latitude was treated as fixed effects, but as with *GS-ELONG*, there was no support for an interaction term.

Study system. To create a genetically diverse ancestral population, we selected 30 accessions of *A. thaliana*, based on published information on glucosinolate profiles (*31*), flowering time (*32*), and trichome density (NASC, The European Stock Centre, www.arabidopsis.info). These lines are maintained in stock centers through selfing and are thus homozygous for most of their genome (*33*). We obtained seeds of all accessions through NASC and propagated them for one generation to amplify seed stocks and reduce potential confounding maternal effects. Three accessions completely failed to germinate; hence we used the 27 remaining accessions for the selection experiment (Table S1). We established laboratory stock cultures of the three aphid species *M. persicae*, *B. brassicae* and *L. erysimi.* Each culture was founded from a single adult aphid which we collected from naturally occurring Brassicaceae in the garden of the University of Zürich at the beginning of the experiment in spring 2009.

Selection experiment. Experimental populations were set up in individual cages made from Plexiglas® (Fig. S6) and were maintained at 18° C under a day/night cycle of 16h/8h. Plant populations in the first generation were established from 20 seeds per genotype (540 seeds in total), and in all subsequent generations 800 seeds were randomly selected to establish new populations in fresh soil. To establish plant populations at the beginning of a generation, seeds were evenly sprinkled into planting trays (18 x 35 cm) filled with 2.5 liters of standard germination soil (pre-mixed with vermiculite, GO M1, Tref Group, The Netherlands). Trays were thoroughly soaked with water, covered with transparent plastic film and cold stratified at 4° C for four days and placed into the cages inside the climate chamber afterwards.

Germination was measured in all populations and 15 days after sowing 50 random plants were sampled per population, removing the fourth leaf from 50 plants and counting the number of trichomes within a defined area on the adaxial surface of the leaf (Ø 4 mm). 17 days after sowing, aphid treatments were initiated by applying six aphids (two per species for aphid mixtures). After the introduction of aphids, cages were checked regularly for cross-contamination among treatments. Contaminations occurred rarely, and only toward the very end of the generational cycle and were therefore considered harmless due to already advanced plant senescence. Since plants in the no-aphid treatment senesced more slowly, contaminations in this treatment were treated with a systemic insecticide (ACTARA®, Syngenta, Switzerland) that killed aphids within a few days. All populations were harvested after 60 days when most plants had senesced. Every generation, the locations of cages in the growth chamber were re-randomized following a modified stratified random design.

Genotyping. We randomly selected 24 plants per population in generation 5 (144 per treatment), and genotyped them using a set of SSLP markers (*34,* see Table S3 for details). Plants were grown in a controlled climate chamber set to 18° C under a day/night cycle of 16h/8h and fifteen days after germination, we harvested two leaves from each plant, placed them into a test tube within a 96-tube rack format, and immediately froze the samples on dry ice for later DNA extraction. Several small glass beads (Ø 1 mm) were added to each tube and 96-tube racks were frozen in liquid nitrogen. Frozen samples were ground by shaking 2 x 30 sec (frequency = 30/sec), turning plates once by 180°, on a Mixer Mill Retsch MM300 (Retsch Technology GmbH, Germany) and DNA was extracted (*35*).

To distinguish the 27 ancestral accessions, we tested five polymorphic loci that produced different fragment size upon amplification (simple sequence length polymorphisms, SSLPs) using PCR (Table S3). We analyzed three markers (*nga6*, *nga172* and *ciw6*) on all plant samples, and a subset of plant samples with ambiguous results was analyzed with two

additional markers (either *nga111* or *ciw3*). PCR was first carried out on DNA of the 27 ancestral lines to establish the reference genotypes. All PCR reactions were carried out using 5 µl DNA and a final reaction volume of 25 µl. The PCR conditions were as follows: 94° C for 120 s (1x); followed by 35 cycles of 94° C for 30 s, 59° C for 20 s, 72° C for 30 s; and 72° C for 10 min (1x). The size of PCR products was analyzed on a QIAxcel® capillary electrophoresis system (Qiagen, Switzerland), using a high-resolution gel cartridge and standard Qiagen reagents. The 15bp/500bp QX Alignment Marker was used to align samples, as PCR product typically ranged between 120 and 250 bp. Following standard Qiagen protocols, we determined DNA fragment size using the BioCalculator Software, which is part of the QIAxcel® system. Plant samples were then assigned to ancestral genotypes using a discriminant analysis (function *lda* in R (*36*)) with fragment sizes of known genotypes as the training sample and unknown plant genotypes as the test sample.

Glucosinolate analysis. We analyzed the chemical profile of an additional set of 24 plants per population of generation 5 (144 plants per treatment). We grew plants in a controlled climate chamber set to 18° C under a day/night cycle of 16h/8h for 20 days, and then we harvested the six largest leaves of each plant. Leaves were placed onto a white plastic surface and photographed for size measurement. Immediately after photographing, all six leaves of one plant were put into a 1.4 ml test tube within a 96-tube rack format (Micronic, The Netherlands), containing 400 µl of 90% methanol, which inhibits the enzymatic breakdown of glucosinolates. Samples were then extracted and analyzed as described in Kliebenstein et al. (*31*). Leaf area of samples was measured from photographs using the open-source image processing software ImageJ (*37*). For a representative subset of genotypes, rosettes were harvested, dried and weighed to generate a calibration line. Using this line, all leaf areas were transformed into leaf masses. Glucosinolate profiles of plant samples were then assigned to

ancestral chemotypes using a discriminant analysis (function *lda* in R) with chemical profiles of known genotypes as training sample and profiles of generation 5 plants as test sample.

Phenotypic screening of *A. thaliana* genotypes. We measured several defense-related traits on the 27 ancestral accessions grown in a controlled climate chamber set to 18° C under a day/night cycle of 16h/8h. These were growth rate (as a descriptor of the competitiveness of a plant (see *25, 38*)), time of flowering, trichome density, and glucosinolate content. Growth rate was measured as biomass growth over the whole plant life using nine sequential harvests. For each harvest, we grew 2-3 individual plants for each of the 27 accessions and harvested plants on days 7, 9, 11, 14, 18, 22, 27, 30 and 38 after sowing. We fitted non-linear growth curves with the function *gnls* implemented in the *nlme* library for R (*36, 39*) to the total above-ground biomass data, using an asymptotic regression and the self-starting routine *SSasymp* (*38*). We calculated size-standardized growth rate (SGR) from the estimated model parameters at a common reference size (mean size 9 days after sowing).

Statistical analyses. All analyses were carried out in R 2.13 for Windows (*36*). For analyses of the selection experiment, aphid treatment was fitted as a five-level factor, with each of the four aphid levels being tested against the no-aphid treatment. All traits with multiple measures per population were analyzed with linear mixed-effects models (lme) implemented in the *nlme* library for R (*39*), using population as a random effect. The change in aphid impact on plant populations over time was analyzed in an lme-model of log-biomass as a function of treatment and generation. In this way, aphid treatment effects are expressed as differences on the log-scale, which is equivalent to log-ratios. Generation was treated as a factor to account for non-linearity in the relation with the response. Average aliphatic and indolic glucosinolate contents in each generation (based on composite leaf samples) were analyzed as log-concentrations in similar lme-models, while for trichome density, the absolute numbers were

analyzed. The overall proportion of the *GS-ELONG* chemotype in generation 5 was analyzed using a generalized linear model (glm) with a binomial error structure.

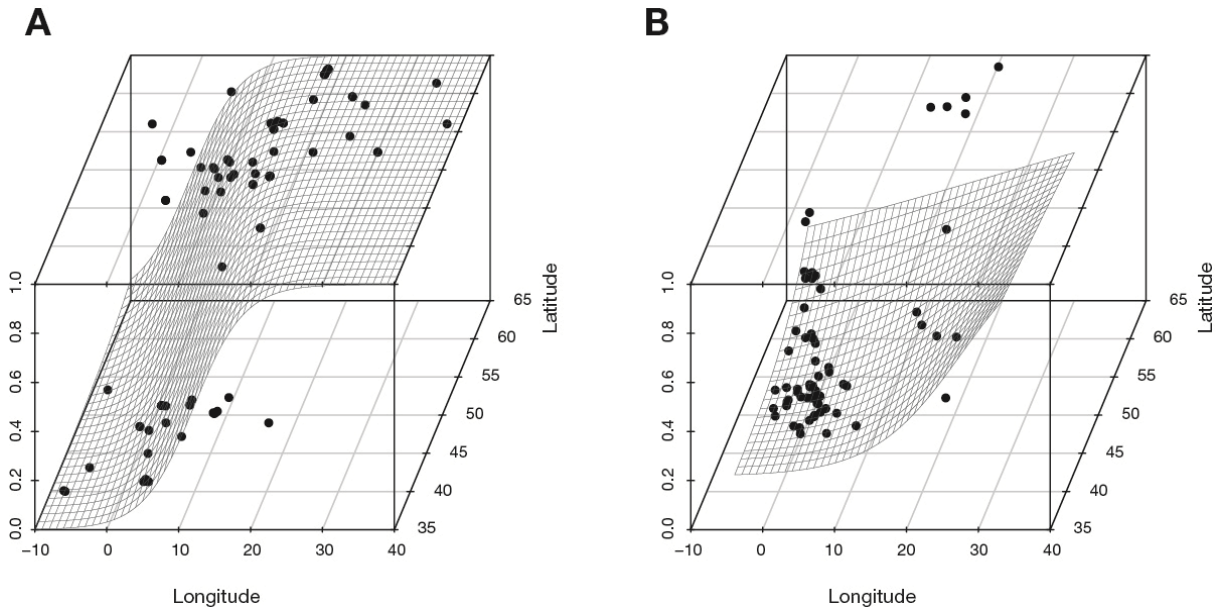


Figure S1. A) Three-dimensional plot of the distribution of *GS-ELONG* across Europe. *4C* chemotypes are represented as zero values and *3C* chemotypes as ones. Each dot represents one accession, and the hatched area is the model fit of a generalized linear model with binomial error structure (binomial glm), *GS-ELONG* as binary response, and longitude and latitude as explanatory variables. **B)** Three-dimensional plot of the relative proportion of the mean number of *Lipaphis erysimi* individuals per *Brevicoryne brassicae* individuals, captured in aerial suction traps distributed across Europe. Each dot represents one trap, and the hatched area is the model fit of a generalized linear mixed-effects model with binomial error structure (binomial glmer); the proportion of the two aphid species from several years of trapping as response, longitude and latitude as explanatory variables, and trap identity as a random effect to account for repeated measures at the same location over multiple years.

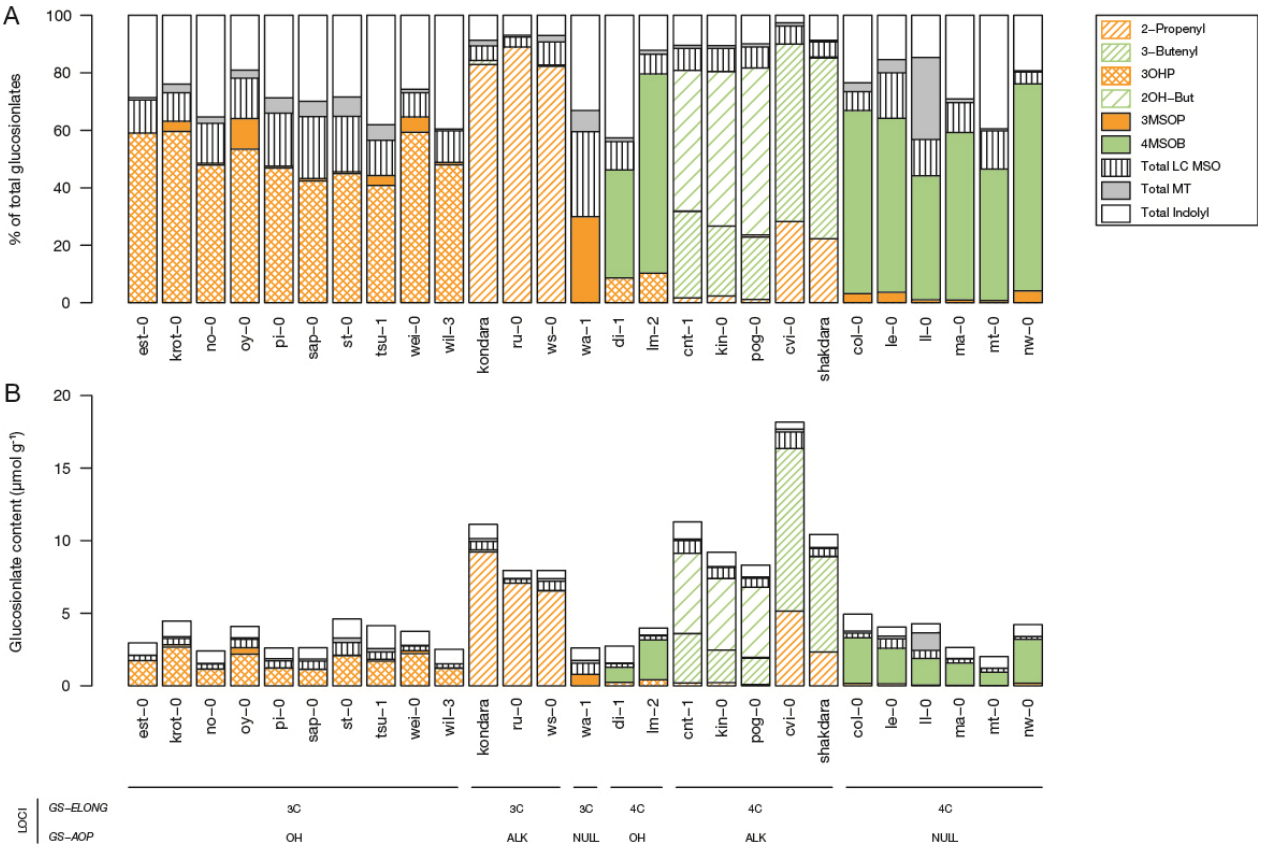


Figure S2. Natural variation in glucosinolate profiles in the 27 accessions (genotypes) of *A. thaliana* used for the selection experiment. Genotypes are ordered according to their chemotype on the basis of allelic variation at two loci, *GS-ELONG* (*3C* and *4C*) and *GS-AOP* (*NULL*, *ALK* and *OH*). Presented are **A)** relative and **B)** absolute concentrations of glucosinolates (µmol g^{-1} dry weight). Compounds with a *3C* side-chain are colored orange, while *4C* compounds are colored green. The biologically active functional group is indicated by the fill of each bar: solid (*NULL*), hatched (*ALK*), and cross-hatched (*OH*). Abbreviations of glucosinolate compounds: *3OHP,* 3-hydroxypropyl; *2OH-But,* 2-hydroxy-3-butenyl; *3MSOP,* 3-methylsulfinylpropyl; *4MSOB,* 4-methylsulfinylbutyl; *Total LC MSO,* long-chain methylsulfinyl; *Total MT,* methylthio.

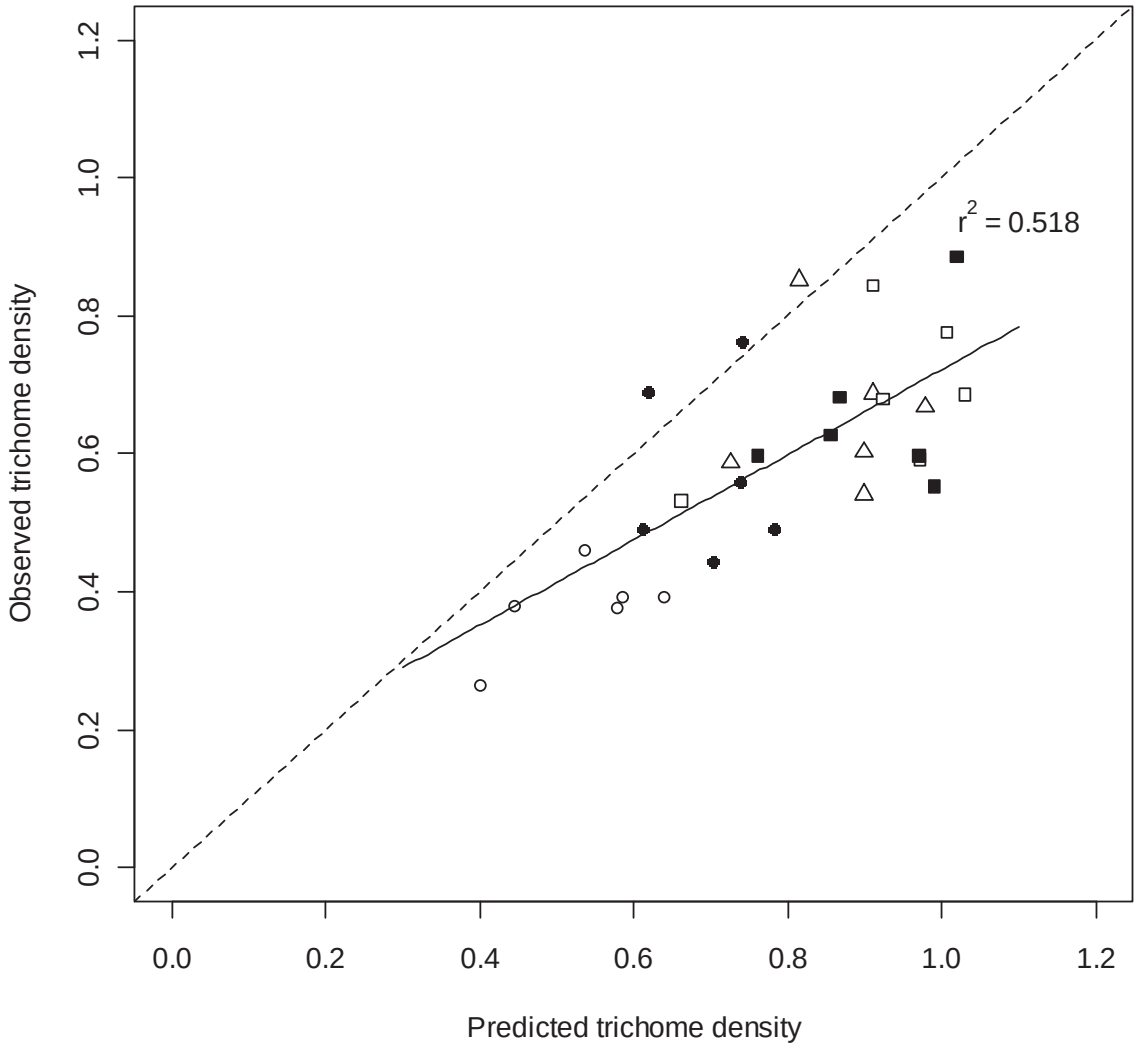


Figure S3. Plot of the predicted *versus* observed trichome density in generation 5 for all 30 populations ($r^2 = 0.52$, $p < 0.001$). Predicted trichome density is based on genotype frequencies in generation 5 and ancestral trait distributions. Symbols represent the five treatments: no aphids (open circles), *M. persicae* (open triangles), *B. brassicae* (filled circles), *L. erysimi* (open squares), and aphid mixture (filled squares). The dotted line indicates the 1:1 line, and the solid line is the actual model fit.

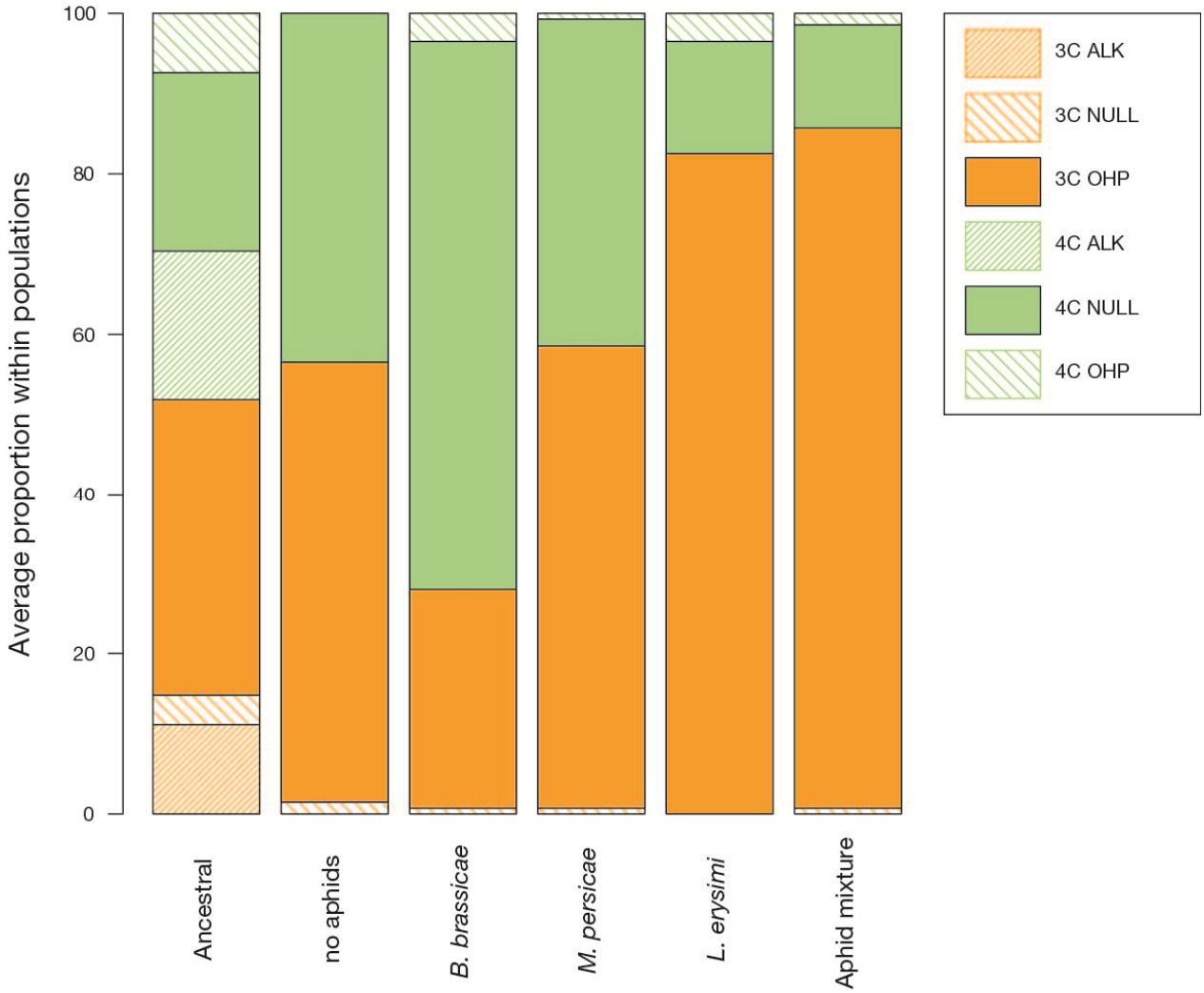


Figure S4. Mean frequencies of the six main chemotypes within plant populations before and after five generations of selection. Mean frequencies are based on HPLC analysis of 24 plants per population (144 plants per treatment). After selection, mainly the chemotypes *3C-OH* and *4C-NULL* remained in all treatments. Both alkenyl chemotypes were lost from all 30 populations, while *3C-NULL* (overall mean proportion: 0.71 %) and *4C-OH* (overall mean proportion: 1.82 %) remained at very low frequencies in all populations.

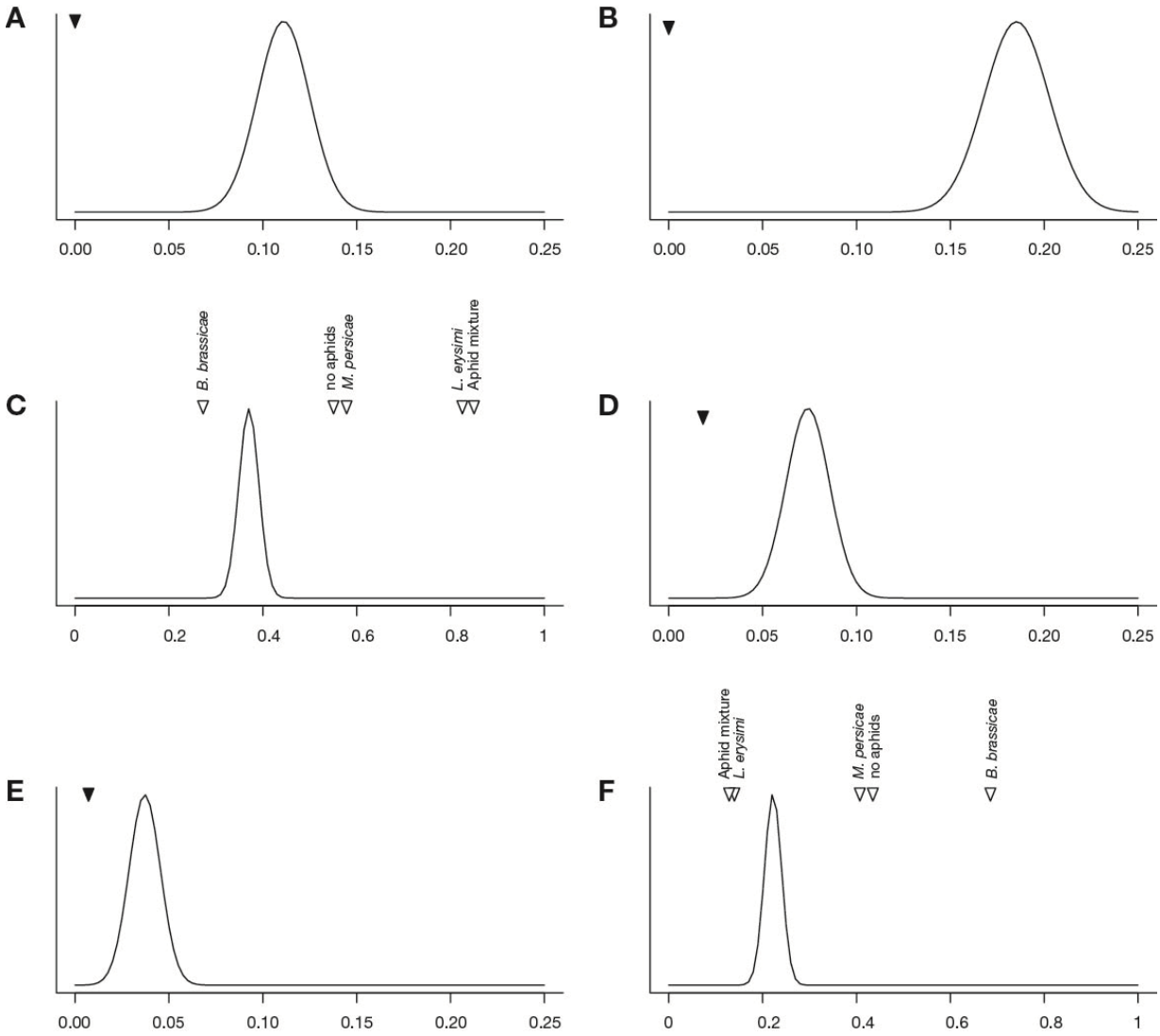


Figure S5. Distribution of expected chemotype frequencies. We generated distributions for the expected chemotype frequencies for **A)** *3C-ALK*, **B)** *4C-ALK*, **C)** *3C-OH*, **D)** *4C-OH*, **E)** *3C-NULL*, and **F)** *4C-NULL*, assuming no treatment differences and only random drift. This was achieved by randomly sampling genotypes based on their frequencies in previous generations only (in generation 1, each genotype had the same probability of being selected). The total population size was constrained to be the total number of seedlings observed in each population in each generation and we include the number of individuals actually genotyped as the final sample. Random sampling was repeated 10'000 times and chemotype frequencies were averaged across all 30 populations. For all observed chemotype frequencies, either the treatment-specific (empty triangles) or the overall average values (filled triangles) are shown, depending on significant treatment differences.

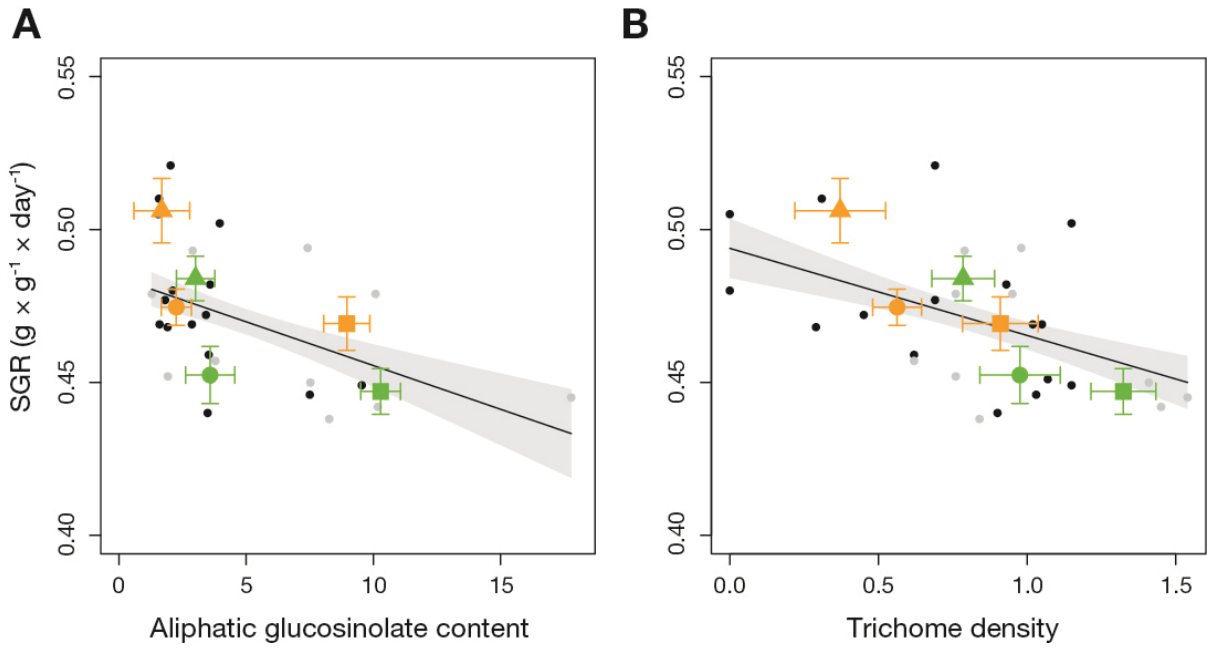


Figure S6. Trait values of the 27 ancestral genotypes. **A)** shows size-standardized growth rate (SGR) and aliphatic glucosinolate content, and **B)** SGR and trichome density. Solid lines are linear regressions with grey areas indicating ± 1 SEs. Genotypes extinct after five generations of selection are colored in grey. The chemotype means ± 1 SEs are overlaid, with symbol shape indicating the *GS-AOP* chemotype (square: *ALK*; circle: *OH*; triangle: *NULL*), and color indicating the *GS-ELONG* chemotype (orange: *3C*; green: *4C*).

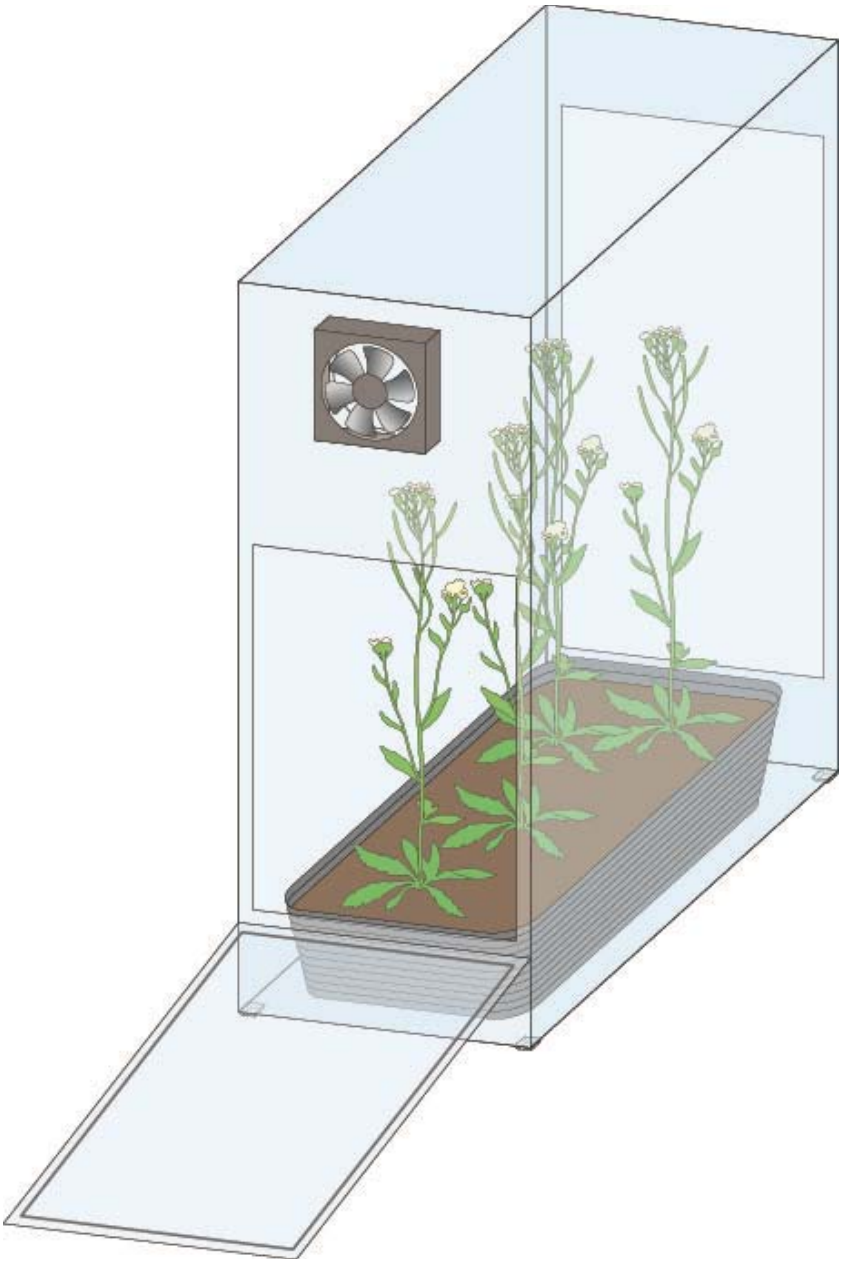


Figure S7. Schematic drawing of an experimental cage. Cages (20 x 36 x 50 cm w x l x h) were used to keep aphid treatments constrained to their respective plant populations in the selection experiment. Cages are made from 5 mm thick Plexiglass® with a square, netted hole in the back (18 x 30 cm) and a flap in the front of the cage for access. The flap is fitted with a rubber seal creating an insect-proof barrier when closed. Each cage is fitted with a ventilator, mounted on top of a circular netted hole (Ø 8 cm) on the front of the cage creating a linear airflow through the cage.

Table S1. List of the 27 genotypes used in the selection experiment. Bolting age and trichome density are mean values from approximately 20 plants. SGR is the average growth rate at a common reference size (mean size at 9 days after sowing). Glucosinolate content is based on the measurement of two plant samples per genotype, each sample consisting of six rosette leaves. All plants were grown in individual pots within a controlled climate chamber set to 18° C under a day/night cycle of 16h/8h.

Genotype	Country of Origin	Line ID	Bolting age [DAG]	Trichomes × $mm^{2\,-1}$	SGR [$g \times g^{-1} \times day^{-1}$]	Chemotype	Aliphatic glucosinolate content [$\mu mol \times g^{-1}$]	Indolic glucosinolate content [$\mu mol \times g^{-1}$]
cnt-1	United Kingdom	N1635	18.5 ± 1.1	1.45 ± 0.18	0.442	*4C-ALK*	10.150	1.184
col-0	United States	N1092	16.3 ± 0.7	1.15 ± 0.25	0.502	*4C-NULL*	3.903	1.155
cvi-0	Cape Verde Islands	N902	18.3 ± 1.0	1.54 ± 0.21	0.445	*4C-ALK*	17.684	0.476
di-1	France	N1108	14.1 ± 0.3	1.05 ± 0.18	0.469	*4C-OH*	1.588	1.169
est-0	Estonia	N1148	14.4 ± 0.7	0.00 ± 0.00	0.480	*3C-OH*	2.106	0.843
kin-0	United States	N1272	18.0 ± 1.7	0.84 ± 0.16	0.438	*4C-ALK*	8.259	0.973
kondara	Tadjikistan	N916	26.1 ± 4.0	0.95 ± 0.12	0.479	*3C-ALK*	10.151	0.966
krot-0	Germany	N3886	17.0 ± 0.8	0.90 ± 0.20	0.440	*3C-OH*	3.396	1.069
le-0	Netherlands	N1308	18.9 ± 1.0	0.62 ± 0.10	0.459	*4C-NULL*	3.490	0.623
ll-0	Spain	N1338	34.3 ± 0.6	0.62 ± 0.13	0.457	*4C-NULL*	3.764	0.628
lm-2	France	N1344	15.7 ± 0.7	1.07 ± 0.19	0.451	*4C-OH*	3.600	0.483
ma-0	Germany	N1356	14.8 ± 0.7	0.69 ± 0.14	0.521	*4C-NULL*	1.904	0.770
mt-0	Libya	N1380	16.1 ± 1.5	0.76 ± 0.12	0.479	*4C-NULL*	1.240	0.790
no-0	Germany	N1394	15.0 ± 1.0	0.69 ± 0.16	0.477	*3C-OH*	1.550	0.846
nw-0	Germany	N1408	16.4 ± 0.5	0.93 ± 0.12	0.482	*4C-NULL*	3.477	0.811
oy-0	Norway	N1643	17.7 ± 0.7	0.45 ± 0.10	0.472	*3C-OH*	3.318	0.776
pi-0	Austria	N1454	19.3 ± 1.0	0.76 ± 0.10	0.452	*3C-OH*	1.867	0.748
pog-0	Canada	N1476	17.9 ± 1.6	1.41 ± 0.17	0.450	*4C-ALK*	7.517	0.823
ru-0	Germany	N1496	16.8 ± 0.4	1.03 ± 0.11	0.446	*3C-ALK*	7.397	0.549
sap-0	Czech Republic	N1506	17.4 ± 1.1	0.29 ± 0.09	0.468	*3C-OH*	1.847	0.787
shakdara	Tadjikistan	N929	19.6 ± 2.2	1.15 ± 0.19	0.449	*4C-ALK*	9.540	0.903
st-0	Sweden	N1534	16.7 ± 1.5	0.79 ± 0.22	0.493	*3C-OH*	3.292	1.306
tsu-1	Japan	N1640	19.0 ± 3.5	0.56 ± 0.12	NA	*3C-OH*	2.568	1.572
wa-1	Poland	N1587	14.2 ± 0.4	0.31 ± 0.13	0.510	*3C-NULL*	1.746	0.866
wei-0	Switzerland	N3110	14.6 ± 0.7	1.02 ± 0.22	0.469	*3C-OH*	2.795	0.966
wil-3	Lithuania	N1598	15.3 ± 0.9	0.00 ± 0.00	0.505	*3C-OH*	1.519	0.994
ws-0	Russia	N1602	NA	0.98 ± 0.20	0.494	*3C-ALK*	7.395	0.557

Values are line means ± 1 SE

Table S2. Statistical tests for change in biomass and trichome densities. All tests are extracted from a linear model of log(Biomass) or a linear mixed-effects model of trichome density as response and treatment, generation, and the interaction as factorial explanatory variables. Values are *t*-tests on the treatment differences from the no-aphid treatment in generation 5.

	log(Biomass)		Trichome density	
M. persicae	$t = 4.16$	$p < 0.001$***	$t = 6.90$	$p < 0.001$***
B. brassicae	$t = 0.97$	$p = 0.333$	$t = 4.69$	$p < 0.001$***
L. erysimi	$t = 2.71$	$p = 0.008$**	$t = 7.36$	$p < 0.001$***
Aphid mixture	$t = 3.78$	$p < 0.001$***	$t = 6.96$	$p < 0.001$***

Table S3. List of primers used for genotyping of the set of 27 *A. thaliana* accessions.

SSLP Locus	SSLP Primer	Chromosome	PCR Primer Sequence	$MgCl_2$ concentration used in PCR reactions
nga111	NGA111F	1	TGTTTTTTAGGACAAATGGCG	1.5 mM MgCl2
	NGA111R		CTCCAGTTGGAAGCTAAAGGG	
ciw3	CIW3F	2	GAAACTCAATGAAATCCACTT	2.5 mM MgCl2
	CIW3R		TGAACTTGTTGTGAGCTTTGA	
nga72	NGA172_F	3	CATCCGAATGCCATTGTTC	2.5 mM MgCl2
	NGA172_R		AGCTGCTTCCTTATAGCGTCC	
nga6	NGA6_F	3	ATGGAGAAGCTTACACTGATC	1.0 mM MgCl2
	NGA6_R		TGGATTTCTTCCTCTCTTCAC	
ciw6	CIW6_F	4	CTCGTAGTGCACTTTCATCA	2.0 mM MgCl2
	CIW6_R		CACATGGTTAGGGAAACAATA	

CHAPTER 6

GENERAL DISCUSSION

GENERAL DISCUSSION

The diversity of life on earth is stunning, yet while this diversity is under threat by human population growth, we still lack the understanding of how it came to be in the first place. Understanding the mechanisms underlying biodiversity is essential if we want to halt the current loss of species. Most species we observe today are part of a complex web of biotic interactions and the loss of these interactions has often unpredictable consequences. Plant-herbivore interactions have received a lot of attention in ecological and evolutionary studies, yet these studies tend to focus on the effects that plant diversity has on communities of herbivores (Elton, 1958; Hutchinson, 1959; Murdoch *et al.*, 1972).

Studies on effects in the opposite direction (i.e., herbivore populations affecting plant diversity) are much rarer, yet it is most likely the interaction of both effects that creates diversity at both trophic levels. Selective pressures exerted by herbivores are responsible for the evolution of a large variety of defensive mechanisms in plants. Defence mechanisms can be plastic traits to some degree, but at the very least still require a genetic basis. Therefore selection for defence traits has the potential to exert selective pressure on genotype or species frequencies. Trade-offs between expression of a trait and plant fitness, as well as between the resistances a trait conveys against different herbivores can then lead to diversifying selection, given a diverse, fluctuating herbivore community. This mechanism of diversity maintenance could be seen as a 'biotic niche concept of diversity', which is the extension of the classical Hutchinsonian niche (Hutchinson, 1961) with added biotic niche axes.

In my thesis, we have tested the two major assumptions underlying the biotic niche concept: defensive traits trade-off with plant fitness, and different herbivore species select for different defence traits. The first assumption is a long-standing prediction in plant defence theory (Coley *et al.*, 1985; Bazzaz *et al.*, 1987; Herms & Mattson, 1992). However, experimental studies have not consistently found these costs (reviewed in Koricheva, 2002).

All chapters of my thesis deal at least in part with the costs of defence. We could clearly and repeatedly show that costs of plant defence exist in *Arabidopis thaliana*, which we used as model plant for all our studies. However, our ability to detect these costs is strongly dependent on the experimental conditions in which plants are grown and the method that are used to describe the costs. We generally grew plants under stressful conditions: we used nutrient-poor soil and we usually added no fertiliser, while in one case, plants in addition experienced high competition. Under such resource-limited conditions allocation costs are most likely to manifest themselves, while all too often especially in genetic studies, plants are grown under ideal conditions and costs are masked as a consequence.

The first four chapters all use size-standardised growth rate (SGR) as the same surrogate measure of fitness. Fitness of any organism is the number of offspring it will achieve in the next generation and in plants, the final lifetime seed production is often considered to be the best estimate of this fitness. However, this is also just a surrogate measure, since seed dispersal, germination and establishment all will further affect the true fitness. Plant growth rate is easier to measure in many plants, and is often a good predictor of the outcome of competition (Fakheran *et al.*, 2010; Züst *et al.*, 2011). Since competition directly affects plant establishment, growth rate can be used as an alternative surrogate measure of fitness. We could repeatedly demonstrate trade-offs between SGR and defensive traits in *A. thaliana*, where genotypes are often either slow growing and highly defended, or fast growing and lacking defences.

Importantly, we found most trade-offs not with total content of chemical compounds, but with a subset of compounds. Glucosinolates, the main chemical defensive compounds of *A. thaliana*, are a highly diverse group of plant metabolites (Benderoth *et al.*, 2006), that most likely evolved in response to different biotic selective pressure. Correlations between SGR and individual glucosinolate compounds were often significantly negative, whereas

correlations with total glucosinolate content were much weaker. When comparing two recombinant inbred line (RIL) populations of *A. thaliana*, we could actually demonstrate that mainly one group of glucosinolates are costly to produce: the aliphatic glucosinolates with a 4-carbon side chain. While these compounds are costly to produce, they also seem to provide the most effective defence against generalist leaf-chewing herbivores. In contrast, many experimental studies working on plant defence only look at the absolute content of toxic chemicals while they neglect the chemical diversity of these compounds. These studies often argue that an herbivore feeding on a plant experiences the whole cocktail of plant secondary metabolites together, and therefore one should not try to compare the relative importance of individual compounds (e.g. Agrawal, 2010).

This argument is only partially valid, since it implies that an individual herbivore can select for the full diversity of plant secondary metabolites within its host plant. Under the second assumption of the biotic niche concept, each herbivore species only selects for one or few chemical compounds, but the full community of different herbivores maintains the chemical diversity within a host plant. In this case, looking at absolute content of toxic chemicals alone would most likely mask most relations between fitness and defence.

In the fifth chapter we found strong evidence for differential selection on plant chemicals by different herbivores. We used three species of aphids and recorded what plant traits were selected in response to constant herbivore pressure over several plant generations. Two of the aphid species are specialised on Brassicaceae plants and therefore are most likely physiologically very similar, yet even these two herbivore species favoured plant populations with very different chemotypes after only five generations. *A. thaliana* is fed upon by a diverse community of herbivores (Mauricio & Rausher, 1997; Bidart-Bouzat & Kliebenstein, 2008), yet not all of these herbivores are present in plant populations at the same time and might be completely absent in some populations in some years. In our selection experiment,

the control treatment without aphids was mimicking this situation. This treatment resulted in the dominance of completely different genotypes, which had generally low levels of defence traits and high growth rates. This confirmed that in the absence of herbivores and any abiotic structure, the most competitive genotypes will outcompete all other, better defended genotypes. Aphid herbivory prevented dominance of these genotypes and favoured different defended genotypes instead. There is extensive evidence in the literature that other types of herbivores favour different plant chemicals (Lambrix *et al.*, 2001; Raybould & Moyes, 2001). In a natural population with fluctuating herbivore communities consisting of aphids, caterpillars, leaf beetles, nematodes, etc., this can therefore explain the high diversity of coexisting genotypes of *A. thaliana*.

We used a relatively artificial model system for both trophic levels. The major deviation from a natural system was the herbivore pressure that occurred in the selection experiment. In the ideal environment of the climate chamber and in the absence of any predators, aphids could reach densities that are never achieved in natural populations. This extreme strength of the selective pressure is also the likely explanation for the strength at which plant populations responded. This does not negate our findings as it still describes a process that occurs naturally at a much reduced strength. Effect sizes which we were able to observe over five generations are thus more likely to happen over several hundred generations in natural populations. The use of aphids as model herbivores is in contrast mainly a conservative choice, since a larger variety on this trophic level would most likely have yielded much stronger differential selection.

A. thaliana is considered mainly a model plant for geneticists (Somerville & Koornneef, 2002) and thus its use for ecological studies is sometimes criticised. However, the vast arsenal of genetic tools available for this plant is a unique resource of which also ecological studies can greatly profit. *A. thaliana* is still a natural plant which has to deal with

the same problems as any other plants: it has to allocate limited resources to grow and reproduce, while it has to fight off competitors and is attacked by pathogens and herbivores. Therefore, our results represent a real ecological process that is most likely happening in most plant species to some degree.

In conclusion, we have found ample evidence that support the biotic niche concept as a mechanism that is at least partially responsible for the genetic and species diversity not only of *A. thaliana*, but most organism at the lowest trophic level. Populations of organisms that are commonly considered as ‘pests’ have thus important ecological functions for the maintenance of this diversity. It is further proof for the fact that habitat protection alone is an insufficient conservation measure to protect diverse ecosystems, if no efforts are made to protect the associated animal and insect communities.

REFERENCES

Agrawal AA. 2010. Current trends in the evolutionary ecology of plant defence. *Functional Ecology* **25**: 420-432.

Bazzaz FA, Chiariello NR, Coley PD, Pitelka LF. 1987. Allocating resources to reproduction and defense. *Bioscience* **37**: 58-67.

Benderoth M, Textor S, Windsor AJ, Mitchell-Olds T, Gershenzon J, Kroymann J. 2006. Positive selection driving diversification in plant secondary metabolism. *Proceedings of the National Academy of Sciences of the United States of America* **103**: 9118-9123.

Bidart-Bouzat MG, Kliebenstein DJ. 2008. Differential levels of insect herbivory in the field associated with genotypic variation in glucosinolates in *Arabidopsis thaliana*. *Journal of Chemical Ecology* **34**: 1026-1037.

Coley PD, Bryant JP, Chapin FS. 1985. Resource availability and plant antiherbivore defense. *Science* **230**: 895-899.

Elton CS. 1958. *The ecology of invasions*. London: Methuen and Co.

Fakheran S, Paul-Victor C, Heichinger C, Schmid B, Grossniklaus U, Turnbull LA. 2010. Adaptation and extinction in experimentally fragmented landscapes. *Proceedings of the National Academy of Sciences of the United States of America* **107**: 19120-19125.

Herms DA, Mattson WJ. 1992. The dilemma of plants: to grow or defend. *Quarterly Review of Biology* **67**: 283-335.

Hutchinson GE. 1959. Homage to Santa Rosalia or why are there so many kinds of animals? *American Naturalist* **93**: 145-159.

Hutchinson GE. 1961. The paradox of the plankton. *American Naturalist* **95**: 137-145.

Koricheva J. 2002. Meta-analysis of sources of variation in fitness costs of plant antiherbivore defenses. *Ecology* **83**: 176-190.

Lambrix V, Reichelt M, Mitchell-Olds T, Kliebenstein DJ, Gershenzon J. 2001. The Arabidopsis epithiospecifier protein promotes the hydrolysis of glucosinolates to nitriles and influences *Trichoplusia ni* herbivory. *Plant Cell* **13**: 2793-2807.

Mauricio R, Rausher MD. 1997. Experimental manipulation of putative selective agents provides evidence for the role of natural enemies in the evolution of plant defense. *Evolution* **51**: 1435-1444.

Murdoch WW, Peterson CH, Evans FC. 1972. Diversity and pattern in plants and insects. *Ecology* **53**: 819-829.

Raybould AF, Moyes CL. 2001. The ecological genetics of aliphatic glucosinolates. *Heredity* **87**: 383-391.

Somerville C, Koornneef M. 2002. Timeline - A fortunate choice: the history of *Arabidopsis* as a model plant. *Nature Reviews Genetics* **3**: 883-889.

Züst T, Joseph B, Shimizu KK, Kliebenstein DJ, Turnbull LA. 2011. Using knockout mutants to reveal the costs of defensive traits. *Proceedings of the Royal Society B: Biological Sciences* **278**: 2598-2603.